

Effective Visual Representations

HABILITATION THESIS

carried out in the years between 2005 to 2015 for the purpose of obtaining the

Venia Docendi (Right to Lecture)

of

Informatik / Computer Science

submitted by

Dr. Ivan Viola

to the Faculty of Informatics
at the TU Wien

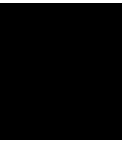
Vienna, 7th September, 2015

Ivan Viola

I would like to express my gratitude to all the researchers I had the pleasure to work with throughout my professional career. You have been standing next to me during all the unforgettable moments, full of curiosity, excitement, and satisfaction. I will never forget these magic moments with you. Thank you!

Contents

1	Overview	1
1.1	Introduction	1
1.2	Illustrative Visualization	2
1.3	Visualization as Applied Perception	5
1.4	Data-Driven Visualization	8
1.5	Conclusions	11
1.6	Authorship Statement	12
	Bibliography	15
2	Papers	19
2.1	A Multidirectional Occlusion Shading Model for Direct Volume Rendering	20
2.2	Illustrative Membrane Clipping	30
2.3	Continuous Levels-of-Detail and Visual Abstraction for Seamless Molecular Visualization	42
2.4	Attractive Flicker: Guiding Attention in Dynamic Narrative Visualizations	56
2.5	A Perceptual-Statistics Shading Model	68
2.6	Perceptually Uniform Motion Space	80



Overview

This thesis presents a selection of my research work in the field of visualization over the span of the last ten years. After obtaining the *Doctor technicae* degree in 2005, I have investigated the adoption of techniques from the visual crafts for interactive data visualization, primarily contributing to the illustrative-visualization subfield. My scientific curiosity was since then centered around the concept of visual abstraction, what it means and how to utilize it for data visualization in various application contexts. This chapter discusses one key aspect of visual abstractions, namely their effectiveness, and it suggests how to systematically design effective visual representations. Publications that form this thesis are in this chapter presented in the light of their design strategy, while their actual main technical contribution is for some of these going beyond the scope of this chapter.

1.1 Introduction

Data visualization is a vehicle that transforms digital information into visual encodings, which convey insights to the humans. The quality of transformation can be characterized according to its expressiveness and effectiveness of providing insight [Mac86]. The expressiveness criteria characterize whether or not all the desired information is communicated. The effectiveness describes how well the visual encoding utilizes the capabilities of the display device and of the human perceptual and cognitive processing. These two characteristics play an important role in the visual mapping stage of the visualization pipeline, where the visual representations are being assigned to the data. While the expressiveness has been an important factor in my research, this thesis focuses on the aspect of effectiveness and how effective visual representations can be obtained.

One can think of a number of systematic strategies to achieve effective visual representations. Here I point out three related but distinct strategies that we investigated over years in numerous research initiatives. The first strategy relates to the adoption of popular designs from the visual communication crafts that were preceding the era of computerized data visualization. The second strategy explicitly takes the knowledge about human visual processing and cognition into account during the design of the visual mapping. The third strategy for designing effective visual representations is based on measurements of the human visual performance in perceptual tasks, and captured into implicit statistical models. In the following, these three strategies will be discussed in further detail, compared, and exemplified on our work that forms this thesis. Especially, I emphasize which opportunities the last mentioned data-driven approach offers. In some sense it represents a fusion of psychophysics and visualization methodologies.

1.2 Illustrative Visualization

Visual communication via illustration or infographics serves various purposes: to convey observations of the visible world or interpretations thereof, to enable classification of objects, to communicate non-visible information by depicting concepts or diagrams, or to convey procedures and processes, to mention a few [VGH⁺05]. The goal of visual communication crafts has been to convey information to humans in a clear and easy-to-understand way. Techniques achieving such a goal were characterized by increased popularity of the employed visual encoding. Considering the development of the visual crafts as an evolutionary process of addressing communication goals, the *surviving species* of visual representations have been favored in the selection process, because they succeeded to communicate the intended information clearer than other techniques. Thus, based on the earlier definition of effectiveness, the effectiveness of traditional visual encodings has controlled the selection process. As a consequence, the visual representations used by visual craftsmen in the past as well as nowadays, should be the most effective ones.

The illustrators have refined their visual designs over centuries. While they often do not include explicit knowledge about human visual perception in the design process, the artists' intuition leads iteratively to visually effective encodings. Illustrators have not been so much concerned with the underlying physiology of human visual processing, when arguing for or against a particular visual encoding. For visual communication craftsmen, some visual designs either *work* well or not, depending on their experience and personal preference [MR77].

For the data visualization pipeline, with effectiveness in mind, utilizing visual encodings from the crafts of illustration or infographics can be the first choice. A substantial body of scientific work has indeed been devoted to mimic visual encodings of traditional visual arts, to analyze these visual encodings, to reverse-engineer their design guidelines, to extend them beyond the realm of a static, non-interactive canvas, and to formalize these into algorithms suitable for computerized image synthesis. The origin of such a scientific endeavor can be dated back to the birth of non-photorealistic graphics at the end of the previous millennium [GG01, SS02]. Later on illustrative rendering techniques have been applied to the visualization of sensed data of medical and biological structures [VKG05, TIP05, BG07], to the visualization of simulations of dynamic flow behaviors [HGH⁺10, BWF⁺10] and complex mechanic interactions [MY⁺10], or to the depiction of abstract non-spatial information [LS07, MM08]. Such approaches are generally known under the term *illustrative visualization*. This thesis contains two works that exemplify visualization techniques inspired from traditional illustration. After presenting these two examples in the following, I will discuss the strengths and weaknesses of illustrative visualization as a strategy for effective visual representations.

Light Placement

The first example, entitled *Multidirectional Occlusion Shading Model for Direct Volume Rendering* [SPBV10], enables real-time computation of shadows using an iterative convolution approach with varying light source positions. Prior work that introduced iterative convolution shadowing for volume rendering [SPH⁺09] is restricted by the constraint that the light vector and the viewing vector need to be identical. When demonstrating the original rendering technique to a professional medical illustrator, she has identified the above constraint as the most severe shortcoming of the technique for attempting to make an effective light staging setup. The light and viewing vector being identical, results in very flat shadows similar to using a flash in photography, where it is mounted on the camera body and flashes directly in the view direction. In an illustration of 3D structures the illustrator would instead virtually illuminate the scene from the top-left or top-right with respect to the viewing direction. Such an illumination setup will give more *depth* to the scene, as formulated by the illustrator, meaning that the depth will be conveyed better. This requirement has prompted us to develop a new technique, which has the same favorable rendering performance characteristics as the prior work, but the constraint of viewing and light direction being

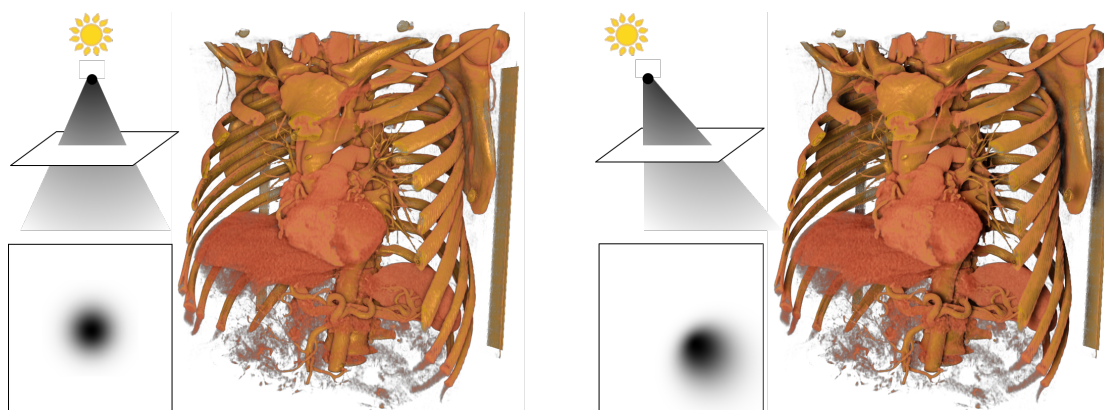


Figure 1.1: Principle and results of iterative convolution-based shadows (left) and (right) our multi-directional occlusion shading approach.

aligned, has been removed. The solution was to use an anisotropic filter kernel of elliptical shape, instead of an isotropic one, for representing the wake of the shadow behind an occluder. The new filter kernel coefficients are obtained by projecting the light cone geometry from a given position onto a slicing plane perpendicular to the viewing direction. Once the anisotropic shadow footprint is calculated, the iterative convolution approach is performed where the occlusion buffer is simply filtered by the anisotropic kernel instead of the isotropic one. The principle and resulting shadowing are depicted in Figure 1.1. The resulting images were able to simulate the top-left or top-right light placement, which was well received by the consulting illustrator.

Clipping

The second example from our work, with the title *Illustrative Membrane Clipping* [BBBV12], was inspired by illustrations of molecular machinery in organisms [Goo98]. Our work builds on the standard clipping technique that is used to show the internal shape of a complex 3D structure, which is for example represented as a volumetric data set. The clipping simply cuts away a portion of the data that is located in one of the two half-spaces defined by the clipping plane. The goal is to show the data located in the other half-space, especially close to the clipping plane interface. Complex structures in the data are however oriented in many different directions, they are in general highly non-planar and clipping can only very rarely show 3D structures in a clear way. Usually, setting a clipping plane results in a compromise, where parts of the interesting structure are revealed through the clipping, but some relevant structures are clipped away. Some less interesting structures might still be present in the visualization and hamper the view on more relevant information.

On the other hand, biological illustrations often show information from the inside of a cell, where clipping geometry is aligned to the structure. The illustrator designs a non-planar cut, so that intended structures are visually conveyed, but less relevant structures are simply not drawn and are *de facto* clipped away. Being inspired by this effective form of illustrative clipping, we have formalized the concept of *membrane clipping*, where the cutting geometry is at the desired cut position and locally adapted to the structures of interest. This is achieved by turning the clipping plane into a deformable *cloth object* simulated in a Newtonian physics solver. An example use case of structure-augmenting slicing is shown in Figure 1.2. A mass-spring system is exposed to a force field which repulses the cloth away from interesting structures, which are then not clipped away and remain visible in their entirety in the resulting visualization.

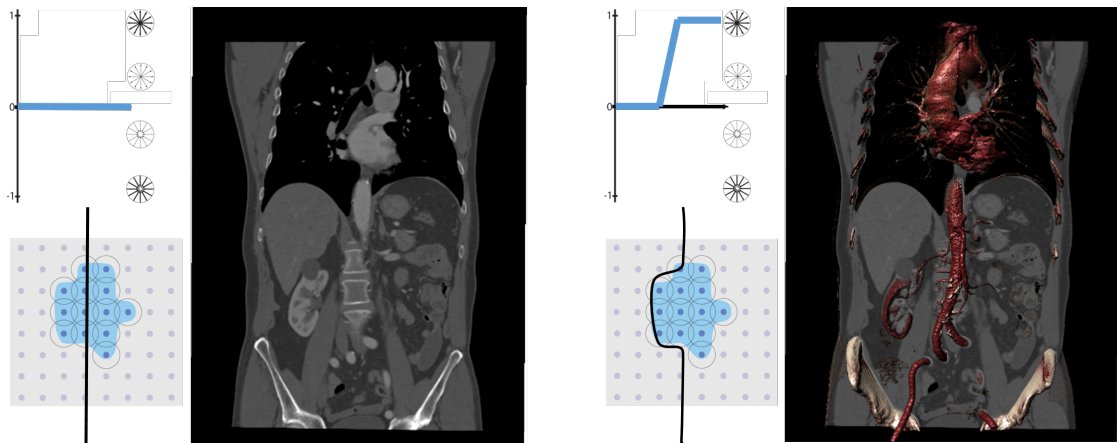


Figure 1.2: Comparison between commonly used planar clipping interaction metaphor (left) and (right) our data-aware membrane clipping approach.

Discussion

Both examples above demonstrate the potential of increasing the effectiveness of visualization by learning from the skills of illustration craftsmen. In the visual communication crafts in general, there is usually no strong need for a scientifically founded argument for or against a particular visual design. Effectiveness of a visual design is first assessed qualitatively by illustrators themselves, as they have training and experiences in how to convey the information clearly and precisely. In the end of the information-supply chain, the effectiveness is assessed through the acceptance by the target audience. Such an indirect form of effectiveness assessment has been to a certain degree propagated to the illustrative visualization techniques that are the computerized counterparts of traditional illustrations. The effectiveness of visual encodings should be ideally demonstrated through an assessment of the human performance in understanding the intended information. The lack of an explicit reference to human visual perceptual processing makes it therefore challenging to define whether and why particular computerized illustrative techniques are effective. It can only be argued that these techniques should be effective, because they are adoptions of illustration techniques that have gained popularity over many years.

Quantification of the effectiveness of a particular illustrative visualization technique is therefore difficult. The question of how effective on a numerical scale or statistically characterized a specific visual encoding is, will in most cases remain unanswered. In practice, the relative effectiveness of a visual encoding has been assessed through a-posteriori comparative, task-oriented user studies. At the end we are left with statements like the illustrative technique A is in terms of correctness or speed by a factor X more effective than other techniques B and C.

Also, the systematic adoption of hand-crafted prior art into interactive visualization has an inherent upper bound in effectiveness, given by the original technique itself. With this adoption strategy it is not readily apparent how to discover new visual encodings that are significantly better than the original technique from illustration.

In summary, the adoption of visual encodings originating from hand-crafted visual designs should be effective. They are probably very good candidates to begin with. In the search for effective visual representations this first strategy can be well argued for. What however might remain unanswered is why it is effective, and how effective the visual encoding is.

1.3 Visualization as Applied Perception

The second approach to achieve effective visual mappings is also the most common in the data visualization research. It is based on the present knowledge from the vision science and psychology that describes how the human visual processing works. Cognitive sciences continuously generate new, evidence-based knowledge about visual processing. The outcome of these sciences is often stated as an argument for a certain visual encoding in the present visualization literature. The century-old Gestalt principles of organization and grouping are a typical example where perceptual psychology theories are used as an argument for a particular visual encoding. Another example is to take into account the knowledge about the duration of perceptual and cognitive functions [CRM91, Joh10], for example how fast or slow a motion of elements on screen should be, to be perceived as motion. Another typical example is the usage of perceptually-uniform color scales used for designing color maps in information visualization, or color scales for dichromats for a particular subset of the population. Visualization researchers would often seek for a textbook on perception [War12, War08, TFCRS11] to test their visualization designs concerning alignment to visual processing. Two works from our research that align the visualization design with outcomes of vision science are presented here, followed by a discussion section.

Object Constancy

The work *Continuous Levels-of-Detail and Visual Abstraction for Seamless Molecular Visualization* [PJR⁺14] describes how the object constancy phenomenon described in the perceptual psychology literature can be utilized in the visual design of a level-of-detail scheme for molecular visualization. The term object constancy refers to the perceived permanence of an object, its shape, color, or size despite of varying viewing and illumination conditions. In our example we utilize the object constancy for a seamless level-of-detail scheme of molecular complexes. With varying levels of detail we are load-balancing the computational effort of rendering molecular surfaces so that the most expensive visual representation is employed for the molecular surface area which is closest to the viewpoint. With increasing eye-space depth we are reducing the computational load by seamless transitions into simplified geometric and illumination representations. The technical concept is based on an implicit surface representation, which can smoothly vary from one geometric representation to another one by linearly interpolating between three-dimensional shapes in object space. This way we vary the geometric complexity from an expensive solvent-excluded surface representation to faster Gaussian blobs. Distant atoms are represented by the van der Waals representation, which is the easiest to compute. Furthermore, at some distance we cluster the atoms into a hierarchical representation and smoothly blend the hierarchical levels. Atoms on an intermediate-node level are grouped into sets of atoms, represented by a larger Gaussian kernel of appropriate size. As the transitions between levels of detail vary smoothly, without any presence of sharp transitions, the object constancy applies. Figure 1.3 illustrates the blending principle and presents visual results of our seamless level-of-detail scheme. The user perceives the simplified representations further away from the viewpoint as being the same type of structural information as the one closer to the viewer. This way the present knowledge about visual perception has shaped our seamless visual abstraction for level-of-detail scheme in the context of molecular visualization.

Visual Guidance

Recently our work on *Attractive Flicker* [WMB⁺14] has utilized the outcome of visual perception research for visual guidance in dynamic scenes on large displays. A typical visual guidance in a static scene that entirely fits into the visual angle of parafoveal vision, can be realized by giving an important object a color distinct from its surroundings. A red object immediately becomes apparent in a scene consisting of blue objects for example. If extending the visual guidance beyond the visual angle of the parafoveal vision, color becomes increasingly ineffective. This also holds for several other visual cues as well and is

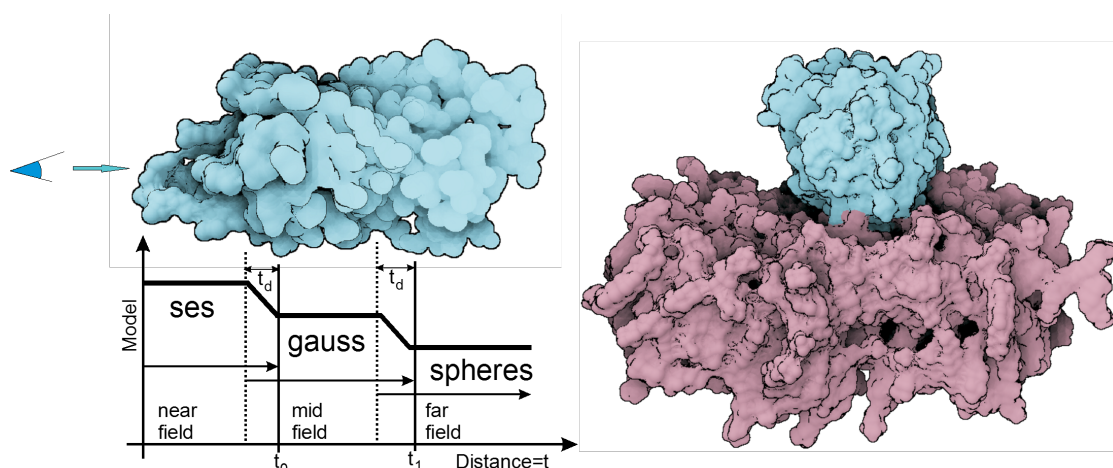


Figure 1.3: Seamless transition in geometric and photometric levels of detail realized as object-space blending of implicit representations. Resulting images show a blend between solvent-excluded surface, Gaussian blobs, and van der Waals sphere objects.

due to color blindness and decreasing visual acuity of the peripheral vision. Peripheral vision, however, is very sensitive to dynamic stimuli. This is true (para-)foveal vision as well, where the stimuli can be motion or other dynamic changes such as flicker in luminance or shape. In a scene, where visual elements are permanently moving, the flicker turns out as the best cue for visual guidance, as it can be easily distinguished from motion. The major problem with visual guidance based on the flicker stimulus is that humans often report strong annoyance, when being exposed to a flickering visual element. In our work we have therefore carefully designed a two-stage guidance signal, where the flicker characteristics are changing over time. In the first stage a strong high amplitude and high frequency sine wave controls the color intensity change, until the guidance brings the object of interest into focus. This stage takes up to one second. The signal is afterwards quickly changing into a small amplitude low frequency intensity pulsation, which is not causing annoyance, but helps the viewer to keep track of the dynamic object of interest. The flicker function is depicted in Figure 1.4. The values of flicker duration, intensity, and frequency have been identified by means of dedicated user studies and were also in agreement with known constants in visual perception from prior literature. Here the knowledge about visual perception helped us to identify a suitable representation for visual guidance concerning dynamic elements on a large display area. Furthermore psychophysics and user research methodologies allowed us to find the proper constants, based on validated thresholds of human visual performance.

Discussion

Visualization of data is designed for a human observer who is perceiving and cognitively processing a given visual representation. Aligning the visual design with the explicit knowledge about human sensory and processing capabilities is a sensible strategy that should lead to effective visual designs. A useful outcome of visual perception research for visualization is a set of explicit guidelines for informed visual design. Then visualization researchers or infographics designers can obey these guidelines making their visual representations effective. This approach is analogous to follow national standards in engineering or architectural design. Without performing a perceptual study, however, the effectiveness of a visual design is not explicitly quantifiable.

The overall task, for which visualization is used, is often quite complex, such as analysis of financial

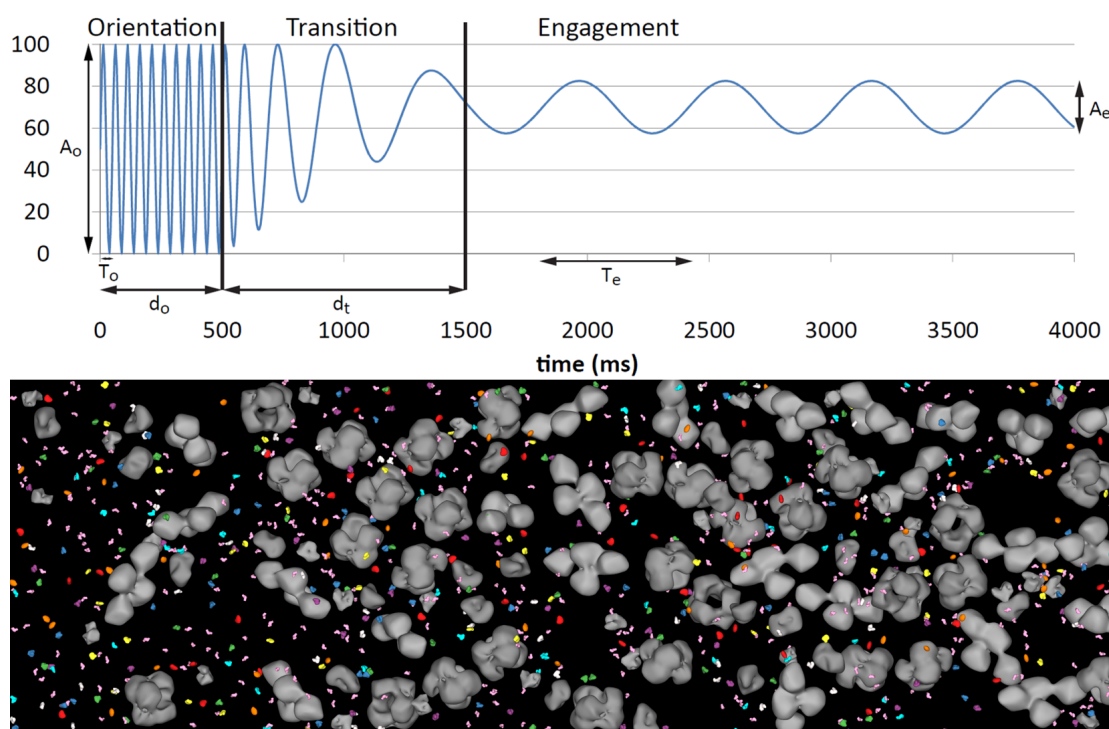


Figure 1.4: The flicker stimulus function starts as a high-frequency strong intensity flicker to attract viewer's attention and smoothly changes into low-frequency intensity pulsation to keep the focus in a dynamic scene. Our visual guidance metaphor is aimed for visualization of dynamic molecular interactions on large displays.

markets, medical diagnosis from multiple imaging modalities, or exploration of seismic and boreholes data in the search for potential oil reservoirs. Visualization employs complex visual abstractions, while vision research probes the visual system with very simple visual stimuli. In the end it becomes quite challenging to assess the effectiveness of a particular visualization, based on the knowledge from the vision sciences. A possible solution is to decompose the complex task into elementary visual aspects such as visual queries, comparisons, or navigations [BGCP11]. These reduced elements of visual encoding are separately assessed in terms of effectiveness from the perceptual point of view. The overall visualization effectiveness can then be evaluated as a weighted sum of its reduced characteristics. Still, the weights can be set arbitrarily, so the overall effectiveness value is hard to assemble.

The assessment of the effectiveness of a visualization, based on the present knowledge of visual processing, creates a directed dependency between the two involved scientific disciplines. Progress in the visualization research becomes dependent on the progress of the vision and psychology sciences. Visualization turns into an applied science of the basic cognitive sciences. If the effectiveness of particular visual variables is not covered by the research in vision sciences, the visualization researcher would first need to conduct perceptual research, before assessing the effectiveness in her specific visualization scenario.

1.4 Data-Driven Visualization

There is at least one more systematic approach for developing effective visualization design. It has the characteristics of a data-driven scientific methodology, where the data are statistics of human perception. Visualization design is tested by an appropriate psychophysics approach that thoroughly probes the efficiency of the design. The outcome of the test are perceptual statistics data from a group of subjects. The perceptual statistics are analyzed for systematic and random error behavior. The visual representation is then modified, based on the outcome of the statistical analysis. The sequence of perceptual probing, analyzing, and redesigning visualizations can be done in several iterations. After several iterations of design improvements, the systematic error falls below a certain value. The resulting visual representation is, statistically-speaking, effective, as the intended information is clearly communicated. While a systematic error describes a bias in perception, a random error describes how uniform the perceptual performance is. In contrast to the two methodologies described in Sections 1.2 and 1.3, this approach results into a quantitative characterization of the effectiveness of the visual mapping under investigation. It does neither strongly depend on the prior work of creative crafts nor on the perceptual sciences.

While it forms a distinct category, the data-driven strategy can be seen as a combination of the previous two strategies. The third category deals with an evolving design, similar to illustrative techniques. The design revision cycles are however much shorter. In illustration the development of new techniques can take decades. In the last category we would have much shorter development revisions. Similar to the second strategy it is also aligned with visual perception performance and it uses the psychophysics methodology to probe this performance. In the second category, the outcome of perception research is in the form of explicit knowledge and is used for informed visualization design. In contrast the third approach bypasses the human interpretation step of the perception-study data. Instead, in an ideal case these data are automatically steering evolution of the visualization design. While the second method would be an informed design, the third method would be a rather blind data-driven design. In the following two examples from my work on data-driven visualization design are presented. The last part will discuss possible future opportunities of this recent design strategy.

Data-Driven Shading Model

In our first work on data-driven visualization design entitled *A Perceptual-Statistics Shading Model* we have chosen a simple scenario, i.e., the visualization of 3D surfaces, which should convey the shape of the structure in the perceptually most correct way [STPV12]. We first visually represented it with a Lambertian diffuse shading model. The correctness of the shape understanding was evaluated by the gauge-figure perceptual task [KvDK92]. We have discovered through a computational analysis that the data contains a systematic error. The eye-space slant of the perceived surface normal is underestimated, as compared to the ground truth. We have derived a very simple correction scheme for the shading method, based on inverting the fitting polynomial curve that characterized the underestimation. In essence we shade the surface using a transformed object's normal with an increased slant, corresponding to the value on the inverted fitted curve. After this correction, we performed another iteration of the user study. This iteration has indicated a statistically-significant improvement for a specific interval of slant angles. The shading correction is depicted in Figure 1.5. Finally, based on the second study, we corrected the shading once again, with the same correction scheme.

Interestingly, the visualization optimization after the first iteration changed the appearance of objects to look like results from the exaggerated shading algorithm [RBD06], only the way of achieving the outcome was fundamentally different. The rendering improvement after the second iteration resembled the light warping method [VPB⁺09], a known non-photorealistic rendering algorithm. The above mentioned prior techniques were inspired by visual arts and by utilizing knowledge about visual perception (Sections 1.2 and 1.3). Our data-driven approach has discovered these visual representations without any prior explicit

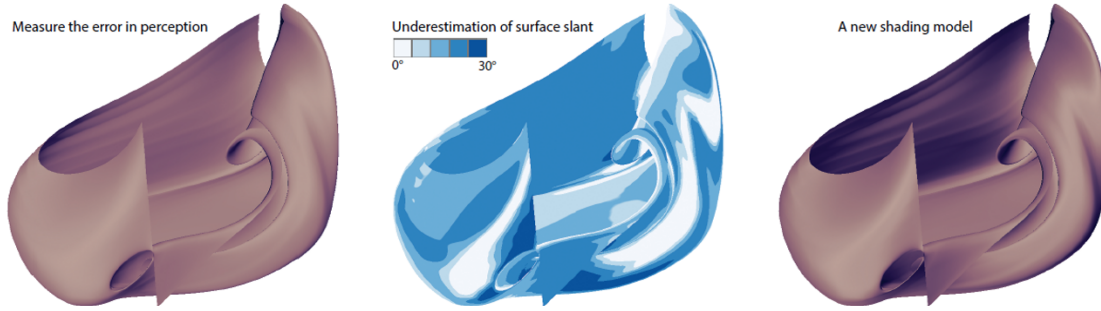


Figure 1.5: The effectiveness of initial shading (left) for conveying surface orientation has been evaluated by means of psychophysics methodology. The systematic error of perceived slant has been isolated (middle). This error forms a basis of the shading correction scheme. Corrected rendering (right) more accurately conveys the surface orientation and is richer on intensity contrast.

knowledge from vision sciences. This observation indicates that the mentioned illustrative techniques have been intuitively optimized for better perception of shape.

Linear Perception of Relative Motion

Our second example of data-driven methodology for designing visualizations is investigating the comparison of motion patterns in the paper *Perceptually Uniform Motion Space* [BTV14]. As an example we start with a 2D flow visualization, where different areas have distinct flow velocities. The research question is to determine whether the humans can guess, based solely on the visualization, how much one flow is faster or slower compared to the reference flow, without providing the numerical value explicitly. We have investigated in which ranges of screen-space speed viewers can easily perceive the difference between two motion patterns, and which multipliers, varying angles and luminance-chromatic contrasts perform best. In a study we asked participants to estimate a multiplication factor on how much the test motion stimulus is faster than the reference one. In this study we have also discovered certain systematic biases. For example if both stimuli were moving rather fast, the users tended to overestimate their difference. With slow motion the difference was rather underestimated. In the case of increasing the speed difference between the reference and the test motion patterns, the difference was systematically underestimated. We devised a scheme for compensating these systematic trends. After the first iteration we show an adapted visual stimulus to the test subjects again and our error metric indicates a constant underestimation, independent of the particular speed settings. After another round of compensation we achieve further improvements in the perceptual test. At the end of our iterative design we obtain a two-dimensional array of motion compensation parameters that correct particle motions so that their speed can be guessed from the visualization without bias. An example of motion patterns and the final compensation model are depicted in Figure 1.6.

Discussion

The investigated data-driven approaches put the goal of the visualization into the main focus, which drives the design evolution until it matches the desired criteria. As such, the strategy is declarative in nature in contrast to the common imperative approaches used nowadays in visualization. Evolutionary declarative design can be applied generally to any visualization type, given that there is a suitable perceptual probing technique, which determines the correctness of understanding the visualization. The initial visualization can start as a result of a computer rendering, but actually it could even be a hand-crafted illustration of

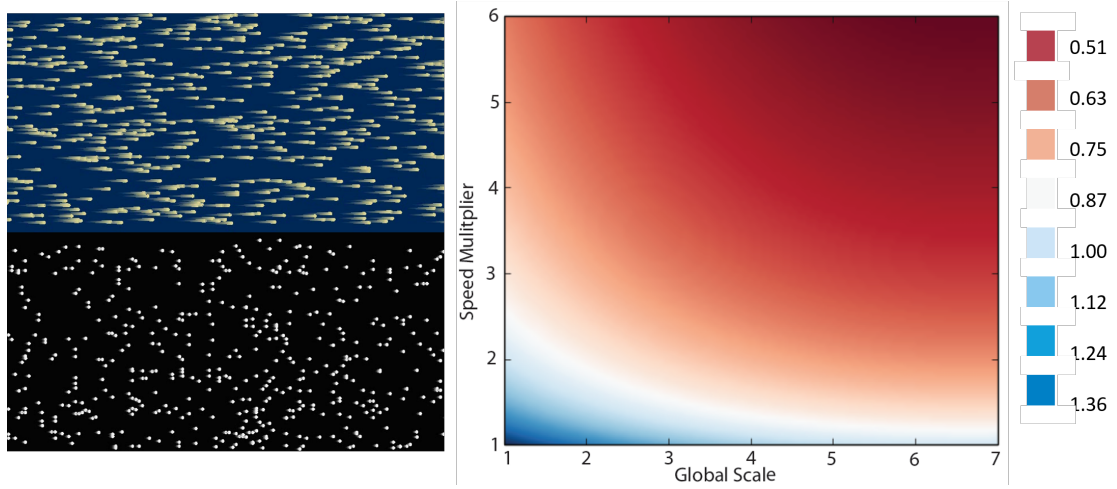


Figure 1.6: Flow visualization stimulus for probing relative motion perception (left) and (right) final motion compensation model. The model describes a magnitude scaling factor, which is a function of overall flow magnitude and the relative speed difference between the reference and test flow patterns.

the data. This initial visual mapping gives an educated guess of how particular data should be visually communicated to the viewer. The efficiency of this visual mapping is evaluated by means of perceptual studies. Then, a statistical analysis will possibly reveal the limitations in the effectiveness by quantitatively measuring the error of the user feedback. Such an error, analogously to the examples above, consists of a random and a systematic error. The improvement of the visualization algorithm is based on the extraction of the systematic error and a corresponding visualization correction strategy. In the above examples we have devised ad-hoc improvement strategies manually. This requires a human involvement in the correction design process.

Such a correction may also be designed automatically by adopting and applying a feedback-loop methodology developed in control theory. Here an open question is how much benefits concepts like *proportional-integral-derivative controllers* and other mechanisms of classical or modern control theory, as well as optimization research, will bring to the visualization design with the human-perception data in the loop.

Once the correction scheme is applied to the initial visual mapping, a perceptual study is conducted, to find out if the systematic error has been reduced. Iteratively, through series of perceptual studies the visualization algorithm converges to its optimal form. Moreover, even if the initial design was achieved by an imperative procedure, every new iteration will become more different from the classical algorithmic visualization approach and will rather be the result of a declarative, data-driven approach.

One could argue that the evolutionary visualization design, where humans are contributing to the design refinement through perceptual studies, is a time-consuming, complex, and expensive process. That is certainly the case and further research should focus on how to accelerate such an evolutionary process. Choosing a suitable optimizer with fast convergence is one possibility. Another possibility is to employ computational perception models in the design process instead of human subjects [PMA⁺15, ZWM13]. These models can be first tested on how they approximate the human visual performance. In case the match is satisfactory, these models can be utilized for intermediate design iterations. The final iteration will be again carried out involving human subjects, to validate the effectiveness estimated by the model. Another option can be to partially re-use the statistics from one iteration in the following iteration to ensure sufficient statistical coverage.

One can even think of a personalized design. This can be achieved by considering samples of one particular person only, instead of analyzing statistics of several people. In such way we could tailor visualizations to particular users, similarly to nowadays writing style or voice recognition approaches. The personalization process itself can be considered by the user as an undesired burden. Therefore, it is important that the perceptual study is part of the user's workflow and collecting statistics does not increase the effort for the user. An alternative is that the perceptual study is carried out in the form of an interactive entertainment and the user enjoys the evolutionary visualization design process.

Once we obtain a sufficient number of samples describing how a particular visualization design has been understood, we can start making predictions on its effectiveness for future use. We can analyze its effectiveness for users with for example different age, gender, social status. The data-driven approach will ultimately make visualization a quantifiable and predictable means of communication.

Besides the listed opportunities, the correction strategy of declarative visualization might lead to some problems. In the shading example the material appearance is changing as a consequence of applying the correction mechanism. This shift in material appearance might in some situations be undesired. In the second example, the speed correction might also lead to undesired artifacts. If for example particles in flow interact with each other, such as they bounce into each other, or cause other forms of events, motion correction might destroy the causal chain for some of these events. Due to the modified speed, the trajectories of two objects will for example not intersect at the same time, but this intersection will happen at different points in time, which can mean that no collision event will happen. Therefore, the form of compensation has to be analyzed on whether it does not cause other severe side effects and problems, while improving the one in focus.

1.5 Conclusions

In this chapter I have reviewed three different strategies for designing effective visualization representations. These methods do not exclude each other. Visualization design might be inspired by an illustration style and simultaneously it can be assessed on how much it is aligned with perceptual guidelines. Such a design could be a starting point for iterative design refinement using the data-driven approach. In this case all three strategies contribute to the visualization design. We have also seen that in case of the perceptual shading model, similar designs can result from totally different strategies. The data-driven approach can also benefit from the fact that data-driven methodology is generally becoming omnipresent in science, so that recently developed and generally applicable machine or deep learning approaches can also be adopted for data-driven visualization. Currently, user interaction, preferences, or personal communication are being analyzed using data science methodologies in order to deliver suitable content and products to the end-user. For the moment we have no good means to provide the user with a personalized visualization that would suit her the best. The data-driven visualization is a promising step in this direction. One can imagine for example an electronic edition of a daily or weekly periodical, where the same underlying information is visually communicated to users with different visualization profiles in a distinct way, on a different level of detail, requiring also a different cognitive load. The publisher will first sample the reader's interest in a particular topic and her level of visualization literacy. Based on these collected data, the user will be given the most detailed visual representation she can cope with for the most interesting topics. The other topics will be represented with higher levels of visual abstraction that are easily accessible. By such a personalization strategy, visualization will gain stronger acceptance among the general population and contribute to a better informed society.

1.6 Authorship Statement

This section gives a short overview on the author's contributions to the publications contained in this thesis. The papers represent a sample of the research work, which the author has published in the years between 2005 to 2015. At the time of writing he has co-authored over 75 peer-reviewed scientific publications, all of them carried out in a collaborative effort together with other researchers. Collaboration on publications is characteristic for the way how our scientific community works, single-author publications occur only in very rare cases. Part of my research has been done at the University of Bergen, Norway, where I worked from 2006-2012 and was finally promoted to a full professor. Another part was done at TU Wien, Austria, from where I graduated and returned back in 2013, after being awarded a Vienna Research Groups grant from the Vienna Science and Technology Fund (WWTF). For all the six selected publications in this thesis I acted as the last author. They all have been published in projects that were granted to me as principal investigator. They are reprinted in this thesis in the form how they were accepted or published, following the publishers policy for reproducing author's work. The publications that constitute this thesis are:

1. A Multidirectional Occlusion Shading Model for Direct Volume Rendering

Veronika Šoltészová, Daniel Patel, Stefan Bruckner, and Ivan Viola; *Computer Graphics Forum*, Volume 29, Issue 3, pages 883–891, 2010

Own contribution: I have developed the initial idea of this paper, which was further investigated and technically realized by my former PhD student Veronika Šoltészová. Daniel Patel and Stefan Bruckner significantly contributed to the idea forming and prototypical development. All authors contributed to the paper writeup led by the first author.

2. Illustrative Membrane Clipping

Åsmund Birkeland, Stefan Bruckner, Andrea Brambilla, and Ivan Viola; *Computer Graphics Forum*, Volume 31, Issue 3pt1, pages 905–914, 2012

Own contribution: I have developed the initial idea of this paper and have coded the basis of the prototype implementation. The technical realization was taken over and completed by my former PhD student Åsmund Birkeland, who has received additional support from Stefan Bruckner and Andrea Brambilla. All authors contributed to the paper writeup led by the first author.

3. Continuous Levels-of-Detail & Visual Abstraction for Seamless Molecular Visualization

Július Parúlek, Daniel Jönsson, Timo Ropinski, Stefan Bruckner, Anders Ynnerman, and Ivan Viola; *Computer Graphics Forum*, Volume 33, Issue 6, pages 276–287, 2014

Own contribution: I have developed the high-level idea on realizing geometric and photometric abstraction as blends between implicit representations. This work has been further expanded by several other contributions and ideas coming from all the other co-authors, most notably from the first author, Július Parúlek who did most of the prototypical development with strong support from Daniel Jönsson. All authors contributed to the paper writeup led by the first author.

4. Attractive Flicker: Guiding Attention in Dynamic Narrative Visualizations

Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, Ivan Viola; *IEEE Transactions on Visualization and Computer Graphics*, Volume 20, Issue 12, pages 2456-2465, 2014

Own contribution: Manuela Waldner has developed the high level idea of this paper. I have contributed with the idea of employing two different signals, one for the peripheral and one for the foveal vision. Mathieu Le Muzic was primarily contributing to the prototype implementation. The experiments have been designed by all authors, primarily by Manuela Waldner and Matthias Bernhard. All authors contributed to the paper writeup led by the first author.

5. A Perceptual-Statistics Shading Model

Veronika Šoltészová, Çağatay Turkey, Mark C. Price, and Ivan Viola; *IEEE Transactions on Visualization and Computer Graphics*, Volume 18, Issue 12, pages 2265-2274, 2012

Own contribution: I have developed the initial idea of this paper, which was taken to the next level by my former PhD student Veronika Šoltészová concerning the further development of the idea, the technical realization, and conducting the perceptual study. Çağatay Turkey was performing the statistical analysis and Mark C. Price has provided his competence in psychological research methodologies. All authors contributed to the paper writeup led by the first author.

6. Perceptually Uniform Motion Space

Åsmund Birkeland, Çağatay Turkey, and Ivan Viola; *IEEE Transactions on Visualization and Computer Graphics*, Volume 20, Issue 11, pages 1542-1554, 2014

Own contribution: I have developed the initial idea of this paper, which was further developed and realized by my former PhD student Åsmund Birkeland. Çağatay Turkey was responsible for the statistical analysis part. All authors contributed to the paper writeup led by the first author.

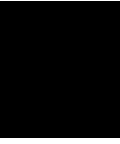
Bibliography

- [BBBV12] Åsmund Birkeland, Stefan Bruckner, Andrea Brambilla, and Ivan Viola. Illustrative membrane clipping. *Computer Graphics Forum (Proceedings of EuroVis)*, 31(3pt1):905–914, 2012.
- [BG07] Stefan Bruckner and M. Eduard Gröller. Style Transfer Functions for Illustrative Volume Rendering. *Computer Graphics Forum (Proceedings of Eurographics)*, 26(3):715–724, 2007.
- [BGCP11] Alexandra Baer, Rocco Gasteiger, Douglas Cunningham, and Bernhard Preim. Perceptual Evaluation of Ghosted View Techniques for the Exploration of Vascular Structures and Embedded Flow. *Computer Graphics Forum*, 30(3):811–820, 2011.
- [BTV14] Åsmund Birkeland, Cagatay Turkay, and Ivan Viola. Perceptually Uniform Motion Space. *IEEE Transactions on Visualization and Computer Graphics*, 20(11):1542–1554, 2014.
- [BWF⁺10] Silvia Born, Alexander Wiebel, Jan Friedrich, Gerik Scheuermann, and Dirk Bartz. Illustrative stream surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1329–1338, 2010.
- [CRM91] Stuart K. Card, George G. Robertson, and Jock D. Mackinlay. The Information Visualizer: An Information Workspace. In *Proceedings of the ACM SIGCHI*, pages 181–186, 1991.
- [GG01] Bruce Gooch and Amy Gooch. *Non-Photorealistic Rendering*. A.K. Peters, 2001.
- [Goo98] David S. Goodsell. *The Machinery of Life*. Springer Verlag, 1998.
- [HGH⁺10] Mathias Hummel, Christoph Garth, Bernd Hamann, Hans Hagen, and Kenneth I. Joy. Iris: Illustrative rendering for integral surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1319–1328, 2010.
- [Joh10] Jeff Johnson. *Designing with the mind in mind: simple guide to understanding user interface design rules*. Morgan Kaufmann Publishers Inc., 2010.
- [KvDK92] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M L Kappers. Surface Perception in Pictures. *Perception & Psychophysics*, 52(5):487–496, 1992.
- [LS07] Martin Luboschik and Heidrun Schumann. Explode to explain - illustrative information visualization. In *Proceedings of Information Visualization*, pages 301–307, 2007.
- [Mac86] Jock Mackinlay. Automating the Design of Graphical Presentations of Relational Information. *ACM Trans. Graph.*, 5(2):110–141, 1986.

- [MM08] Kevin McDonnell and Klaus Mueller. Illustrative parallel coordinates. *Computer Graphics Forum*, 27(3):1031–1038, 2008.
- [MR77] Michael Macdonald-Ross. How numbers are shown. *AV Communication Review*, 25(4):359–409, 1977.
- [MY⁺10] Niloy J. Mitra, Yong-Liang Yang, Dong-Ming Yan, Wilmot Li, and Maneesh Agrawala. Illustrating how mechanical assemblies work. *ACM Transactions on Graphics*, 29(4):58:1–58:12, 2010.
- [PJR⁺14] Julius Parulek, Daniel Jönsson, Timo Ropinski, Stefan Bruckner, Anders Ynnerman, and Ivan Viola. Continuous levels-of-detail and visual abstraction for seamless molecular visualization. *Computer Graphics Forum*, 33(6):276–287, 2014.
- [PMA⁺15] Steven Poulakos, Rafael Monroy, Tunc Aydin, Oliver Wang, Aljoscha Smolic, and Markus Gross. A computational model for perception of stereoscopic window violations. In *Quality of Multimedia Experience (QoMEX), 2015 Seventh International Workshop on*, pages 1–6, 2015.
- [RBD06] Szymon Rusinkiewicz, Michael Burns, and Doug DeCarlo. Exaggerated Shading for Depicting Shape and Detail. *ACM Transactions on Graphics*, 25(3):1199, 2006.
- [SPBV10] Veronika Soltészová, Daniel Patel, Stefan Bruckner, and Ivan Viola. A multidirectional occlusion shading model for direct volume rendering. *Computer Graphics Forum (Proceedings of EuroVis)*, 29(3):883–891, 2010.
- [SPH⁺09] Mathias Schott, Vincent Pegoraro, Charles Hansen, Kévin Boulanger, and Kadi Bouatouch. A directional occlusion shading model for interactive direct volume rendering. *Computer Graphics Forum (Proceedings of EuroVis)*, 28(3):855–862, 2009.
- [SS02] Thomas Strothotte and Stefan Schlechtweg. *Non-photorealistic Computer Graphics: Modeling, Rendering, and Animation*. Morgan Kaufmann Publishers Inc., 2002.
- [STPV12] Veronika Soltészová, Cagatay Turkay, Mark Price, and Ivan Viola. A perceptual-statistics shading model. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2265–2274, 2012.
- [TFCRS11] William Thompson, Roland Fleming, Sarah Creem-Regehr, and Jeanine Kelly Stefanucci. *Visual Perception from a Computer Graphics Perspective*. A. K. Peters, Ltd., 1st edition, 2011.
- [TIP05] Christian Tietjen, Tobias Isenberg, and Bernhard Preim. Combining silhouettes, surface, and volume rendering for surgery education and planning. In *Proceedings of EuroVis*, pages 303–310, 2005.
- [VGH⁺05] Ivan Viola, M. Eduard Gröller, Markus Hadwiger, Katja Bühler, Bernhard Preim, and David Ebert. Illustrative Visualization. In *Eurographics 2005 - Tutorials*, pages 187–329, 2005.
- [VKG05] Ivan Viola, Armin Kanitsar, and Meister Eduard Gröller. Importance-driven feature enhancement in volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):408–418, 2005.
- [VPB⁺09] Romain Vergne, Romain Pacanowski, Pascal Barla, Xavier Granier, and Christophe Schlick. Light Warping for Enhanced Surface Depiction. *ACM Transactions on Graphics*, 28(3):1, 2009.

- [War08] Colin Ware. *Visual Thinking: For Design*. Morgan Kaufmann Publishers Inc., 2008.
- [War12] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., 3 edition, 2012.
- [WMB⁺14] Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, and Ivan Viola. Attractive flicker: Guiding attention in dynamic narrative visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2456–2465, 2014.
- [ZWM13] Lin Zheng, Yingcai Wu, and Kwan-Liu Ma. Perceptually-based depth-ordering enhancement for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):446–459, 2013.

CHAPTER 2



Papers

2.1 A Multidirectional Occlusion Shading Model for Direct Volume Rendering

Veronika Šoltészová, Daniel Patel, Stefan Bruckner, and Ivan Viola

Computer Graphics Forum, Volume 29, Issue 3, pages 883–891, 2010

DOI: 10.1111/j.1467-8659.2009.01695.x

The definitive version is available at <http://diglib.eg.org/>
and <http://onlinelibrary.wiley.com/>.

A Multidirectional Occlusion Shading Model for Direct Volume Rendering

Veronika Šoltészová¹ Daniel Patel² Stefan Bruckner³ Ivan Viola¹

¹University of Bergen, Norway

²Christian Michelsen Research, Norway

³Simon Fraser University, Canada

Abstract

In this paper, we present a novel technique which simulates directional light-scattering for more realistic interactive visualization of volume data. Our method extends the recent directional occlusion shading model by enabling light-source positioning with practically no performance penalty. Light transport is approximated using a tilted cone-shaped function which leaves elliptic footprints in the opacity buffer during slice-based volume rendering. We perform an incremental blurring operation on the opacity buffer for each slice in front-to-back order. This buffer is then used to define the degree of occlusion for the subsequent slice. Our method is capable of generating high-quality soft shadowing effects, allows interactive modification of all illumination and rendering parameters, and requires no pre-computation.

Categories and Subject Descriptors (according to ACM CCS):

I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism —Color, shading, shadowing, and texture

1. Introduction

Local illumination models, such as the Phong model, are suited for conveying shape cues for well-defined structures in volume data. However, they are generally not suitable for visualization when the main goal is to emphasize three-dimensional structural arrangements. In such a scenario, it is important to convey information about relative positions and distances between individual features. The human visual system is adapted to inferring three-dimensional information from illumination. Soft shadows, in particular, are effective monocular depth cues. Not only do they provide occlusion information, but the size and shape of the penumbra can be used to judge distances. Global illumination models provide these cues at high computational costs, especially for volume rendering. Visualization research has therefore focused on the development of new global illumination approximations for volume data that limit the complexity and allow for real-time image synthesis. For this purpose, precomputation or parameter constraint strategies are frequently employed. Both suffer from limited flexibility which can be problematic when interactive exploration is required. The directional oc-

clusion shading model introduced by Schott et al. [SPH*09] is a forward scattering approximation based on a conical phase function. While the method is capable of generating realistic illumination at interactive frame rates, it requires that the view and the light directions have to coincide. In this paper, we introduce a multidirectional occlusion model, which removes this constraint.

The importance of illumination in 3D object perception has been well-studied [BBC83,KMK94,BLK00]. To find out how to best improve volume rendering we have been conducting studies with medical illustrators. During our demonstrations of state-of-the-art visualization techniques to experienced medical illustrators, their first critique point was the positioning of the light in the scene and the choice of non-standard colors. While visualization researchers often carelessly define the light vector parallel to the view vector, this is considered a novice mistake in the domain of illustration. The resulting image is *flat*, akin to photos taken with built-in front flash. To give depth to an image, as a rule, medical illustrators use illumination from the top left. To further op-

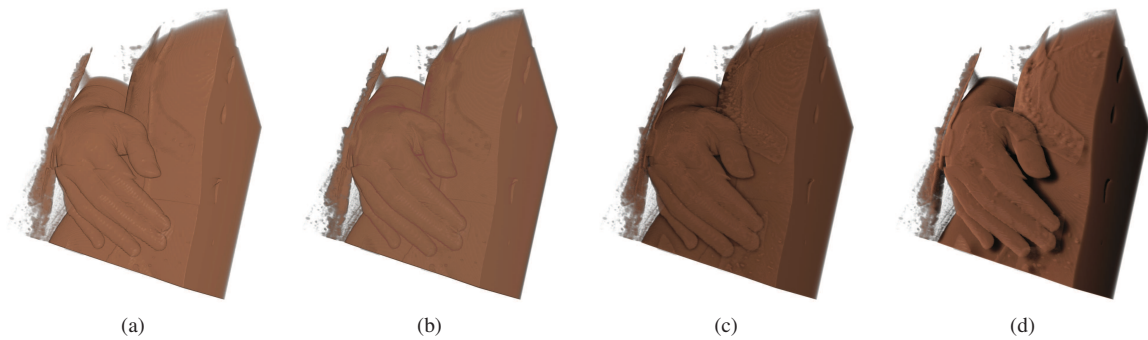


Figure 1: Visualizations of a human hand using raycasting (a), and sliced-based volume rendering (b), both using the Phong illumination. Directional occlusion shading model with a headlamp illumination setup (c) and illumination from top-left (d). Illumination with top-left light position causes that the fingers cast soft shadows on the body and evoke strong depth-perception cues.

timize the appearance of the depicted structures, manual fine tuning is required.

The directional occlusion shading model for interactive direct volume rendering takes the advantage of a headlight setup for performance reasons: by placing the light source at the eye position, the samples required for illumination can be reused for compositing, allowing the method to perform both operations in a single pass for a view-aligned slice through the volume. Our approach uses elliptic occlusion footprints computed from the light position, instead of the symmetric spherical footprints which were used in the original paper. We achieve the same performance with the additional possibility to position the light source anywhere within the hemisphere defined by the view vector. An example of the significant improvement of depth perception compared to the previous method is shown in Figure 1. A visualization of a human hand rendered with different techniques is compared to the headlight and top-left shading. Both a professional illustrator and a user study confirmed our subjective assessment which favored the rendering result 1(d).

The remainder of this paper is structured as follows: In Section 2 we review related work. Our multidirectional occlusion model is derived in Section 3. Section 4 provides additional implementation details. Results are presented and discussed in Section 5. Conclusions are drawn in Section 6.

2. Related Work

Levoy [Lev87] proposed the use of gradients in volume rendering for evaluating a surface-based local illumination model. While this common approach is effective in accentuating material boundaries, it suffers from noise. In particular, gradient-based shading fails to provide useful results in nearly homogenous regions. Illumination models which exploit the volumetric nature of the data can therefore pro-

vide additional cues. Max [Max95] gives a comprehensive overview of different optical models for volume rendering.

Yagel et al. [YKZ91] employed recursive ray tracing which allows for effects such as specular reflection and shadows. Behrens and Ratering [BR98] added shadows to texture-based volume rendering by using an additional shadow volume. The model presented by Kniss et al. [KKH02, KPH*03] captures volumetric light attenuation effects including volumetric shadows, phase functions, forward scattering, and chromatic attenuation using half-angle slicing. Hadwiger et al. [HKS06] presented a GPU-accelerated algorithm for computing deep shadow maps for volume rendering. Rezk-Salama [RS07] proposed a semi-interactive approach for GPU-based Monte Carlo volume raytracing.

Ambient occlusion as described by Zhukov et al. [ZIK98] inverts the principle of light-exposure of a point in space to obscurity by its close environment. Dachsbacher et al. [DSDD07] refer to obscurity as antiradiance. They treat visibility implicitly while propagating antiradiance as an additional quantity. The advantage of these approaches is that they are view-independent: for fixed geometry, occlusion information only has to be computed once and can then be applied efficiently during rendering, for example using texture mapping. Several fast techniques which utilize this concept have been presented [Bun05, SA07]. Knecht [Kne07] and Méndez-Feliu [MS09] provide comprehensive overviews of rendering techniques based on ambient occlusion and obscurities.

In the context of volume visualization, the radiance at a point is determined by shooting rays in all directions from the point and averaging its degree of occlusion by other parts of the volume. The result is an approximation of global diffuse illumination. It produces soft shadowing effects which give a good indication of spatial relationships. However, the

opacity at any point is determined by the transfer function. Ambient occlusion therefore requires an expensive computation step every time the transfer function is modified. Stewart [Ste03] introduced vicinity shading, a variation of ambient occlusion to enhance perception of volume data by darkening depressions and crevices. To reduce evaluation costs, occlusion calculations are reused. The approach of Ropinski et al. [RMSD*08] relied on local histogram clustering to precompute occlusion information for all possible transfer function settings. However, high frequency data, in particular the presence of noise, reduces the effectiveness of their clustering approach and can lead to artifacts. Additionally, their precomputation process is very time and memory consuming. Hernell et al. [HYL07] used a local approximation of ambient occlusion in volumes to limit computation times. In subsequent work [HLY08, HLY09], they utilized local piecewise integration to approximate global light propagation. This approach still requires ambience data for each voxel to be recomputed when changing the transfer function, but their method is able to run interactively by limiting the number of rays shot for evaluating the ambience and by subsampling the rays using adaptive compression. In recent work, Ropinski et al. [RDR10] described a volumetric lighting model which simulates scattering and shadowing. They use slice-based volume rendering from the view of the light source to calculate a light volume and raycasting to render the final image.

View-dependent approaches do not require extensive precomputation and therefore allow fully interactive transfer function modification. This is frequently achieved by limiting light evaluation from spherical local neighborhoods to conical neighborhoods. Desgranges et al. [DEP05] use incremental blurring to achieve shading effects without the use of a gradient. The approach by Bruckner and Gröller [BG07] employed non-photorealistic shadowing and emission effects for the purpose of illustration. Finally, as stated in the previous section, our method is an extension of the model by Schott [SPH*09].

3. Multidirectional Occlusion Shading

The Directional Occlusion Shading model by Mathias Schott et al. (MS-DOS) [SPH*09] describes an approximation of light-scattering in particles of a volume. This simple method generates soft a shadow effect and hence provides important shape and depth-perception cues. Although the approximation of the light transfer delivers slightly different results compared to reference images from a raytracer, it provides visually compelling shading effects at interactive frame-rates and with no precomputation. However, the light transfer approximation in the MS-DOS model constrains the light direction to the viewing direction. In this section we derive an approximation which does not limit the light to this fixed direction.

3.1. Light Transport Equation

The directional occlusion shading model approximates transport of light energy L in a medium. Every point in the environment receives a portion of energy, i.e., radiance composed by background radiance L_b and medium radiance L_m . The medium radiance consists of the emitted radiance L_e and in-scattered radiance L_i . The emitted radiance at a point \mathbf{x} depends only on the local environment of \mathbf{x} . Unlike L_e , the in-scattered radiance L_i integrates over global features:

$$L_i(\mathbf{x}, \omega) = \int_{4\pi} L(\mathbf{x}, \omega_i) \Phi(\omega, \omega_i) d\omega_i \quad (1)$$

where $\Phi(\omega, \omega_i)$ denotes the phase function for two light-ray directions ω and ω_i . L_i quantifies the total radiance incident to point \mathbf{x} from all directions ω_i . From Equation 1, it can be seen that L_i requires an expensive recursive evaluation. The MS-DOS shading model and our model (multidirectional OS) simplify the evaluation which considerably reduces the computational costs.

We assume that the medium emits light only in directions within a specific cone. The phase function from Equation 1 can be therefore replaced by a simple cone-shaped phase function $\Phi_{\theta, \alpha}(\omega, \omega_i)$ where θ is the aperture angle and α the tilt angle of the cone. A schematic illustration of this scenario is depicted in Figure 2. A particle at a point \mathbf{x} scatters light which is received by particles inside the cone. The in-scattering term L_i is conceptually related to the fractional visibility which is equivalent to the opacity and cumulates information about ambient occlusion.

Like the original model, we use a slice-based volume renderer with an additional opacity buffer. Slices are composited in the front-to-back order and the opacity buffer is incrementally filtered and used to determine the accumulated opacity for the next slice as shown in Figure 3. MS-DOS

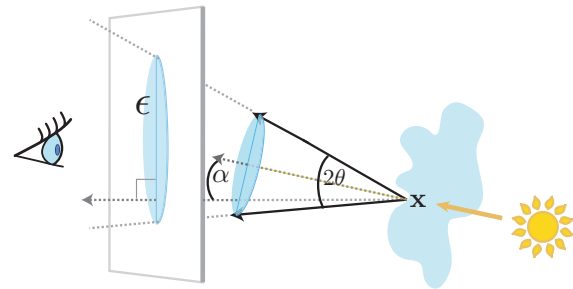


Figure 2: Conical phase function setup: a selected point in space \mathbf{x} scatters light which we approximate by a tilted cone (α = tilt, θ = aperture). The axis of the cone is parallel to the light direction. The projection of the light energy leaves an elliptical footprint ϵ on a selected viewing plane.

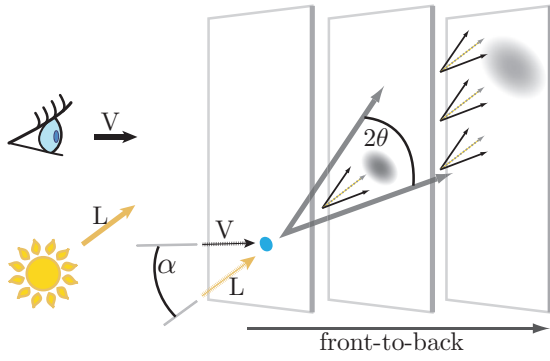


Figure 3: Incremental blurring of the opacity buffer. We use a view-aligned slice-stack composited in the front-to-back order.

operates on view-aligned slices and assumes that the direction of the light source is aligned to the viewing direction. As a consequence, the opacity buffer can be convolved with a symmetrical disc-shaped Gaussian kernel. To enable interaction with the light source, we change the symmetrical disc-shaped kernel to an elliptical kernel. The ellipse ϵ is defined by the intersection of a tilted cone which represents the phase function $\Phi_{\theta, \alpha}(\omega, \omega_i)$ and the slice plane. The cone-shaped phase function is tilted by an angle α which is limited to $[0, \frac{\pi}{2} - \theta)$. This restricts the cone-section from degenerating into hyperbolas or parabolas. Figure 4 describes this geometrical situation.

3.2. Analysis of the Geometrical Model

This section describes the analytical computation of the elliptical kernel, namely, the derivation of its major and minor axes $A = |A_1A_2|$ and $B = |CC'|$ from a known tilt α , a cone aperture θ and a slice distance d . According to Figure 4(a), we derive R from d , θ and α as:

$$R = d \frac{\tan \theta}{\cos \alpha} \quad (2)$$

The axis of the cone intersects the plane at the point O . When the tilt angle $\alpha = 0$, the cone section is a circle, and $a_1 = a_2 = A$. With a known R , we turn to the law of sine in the triangles $\triangle A_1V_1O$ and $\triangle OA_2V_2$. With α , θ , and R given, Equations 3 and 4 yield a_1 and a_2 :

$$\frac{a_1}{\sin(\frac{\pi}{2} - \theta)} = \frac{R}{\sin(\frac{\pi}{2} + \theta - \alpha)} \quad (3)$$

$$\frac{a_2}{\sin(\frac{\pi}{2} + \theta)} = \frac{R}{\sin(\frac{\pi}{2} - \theta - \alpha)} \quad (4)$$

$$A = \frac{a_1 + a_2}{2} \quad (5)$$

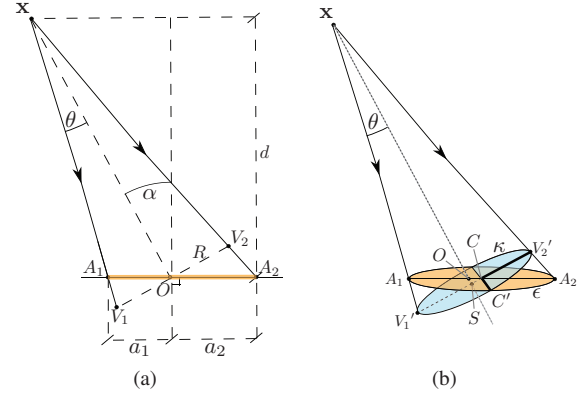


Figure 4: A geometrical description of the cone-shaped phase function: the elliptical cone-section defines a circle κ intersecting the center C of the ellipse ϵ . A side view (a) and a 3D view (b) of the planar cone-section.

With known a_1 and a_2 , we use Equation 5 to calculate A which is the major axis of the ellipse.

The center of the ellipse ϵ is in C with $\|OC\| = \|\frac{a_2 - a_1}{2}\|$. We define a circular cone section κ which intersects the point C . Consequently, the axis of the cone intersects κ in its center S . This scenario is illustrated in Figure 4(b). The intersection line $\epsilon \cap \kappa$ is perpendicular to A_1A_2 and intersects the center C of ϵ . Consequently, $\epsilon \cap \kappa$ is collinear with the minor axis of ϵ . Figure 5 illustrates the side view of $\epsilon \cap \kappa$. In Figure 6, we focus on the triangles $\triangle XV_2V_2$ and $\triangle SCO$, and on the circle κ . Basic analysis implies Equations 6, 7, 8, and 9. Solving them, we determine B - the minor axis of ϵ .

$$\sin \alpha = \frac{d'}{\|OC\|} \quad (6)$$

$$\tan \theta = \frac{dR}{d'} \quad (7)$$

$$R' = R + dR \quad (8)$$

$$B = \sqrt{R'^2 - \|OC\|^2 + d'^2} \quad (9)$$

3.3. Weighting Function

Light rays collinear to the cone axis hit the slice with the highest intensity. We revisit Figure 4(b): O is the point with the highest incident energy. We define a weighting function as follows:

$$W_L(x, y) = 1 - k \quad (10)$$

with k defined implicitly by:

$$\frac{(x - (1 - k)\|OC\|)^2}{A^2} + \frac{y^2}{B^2} = k^2 \quad (11)$$

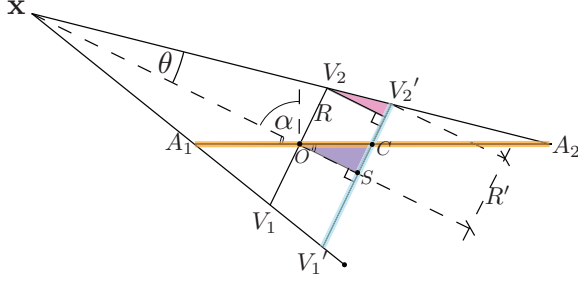


Figure 5: A detailed side-view of the intersection of ellipse ϵ and circle κ .

Equation 11 results in a quadratic equation with two real roots from which we take the maximum. A kernel with a linear fall-off from O towards the border of the ellipse is illustrated in Figure 7(a). Additionally, we apply the Gaussian function to smooth the fall-off of the weights as illustrated in Figure 7(b).

3.4. On-the-fly Convolution

We apply an incremental convolution of the opacity buffer O_i and the elliptical kernel G_ϵ for each slice i . As the light direction \mathbf{L} changes, G_ϵ has to be aligned respectively. We project the light vector to the viewing plane which yields a 2D-vector \mathbf{L}' and rotate the kernel so that its major axis is aligned with \mathbf{L}' :

$$\frac{\vec{OC}}{\|\vec{OC}\|} = \frac{\mathbf{L}'}{\|\mathbf{L}'\|} \quad (12)$$

The weight-distribution in G_ϵ depends only on the tilt, aperture, light direction, and slice distance. Therefore, an update is triggered only if one of these parameters changes. In practice, we render the kernel G_ϵ to a texture when an update is triggered. First, we uniformly scale the ellipse so that it fits into a unit-square. Second, we set-up the texture coordinates so that G_ϵ is aligned correctly. During the volume rendering pass, we apply inverse scaling operation to regenerate G_ϵ of the correct size. In Figure 8, we visualize the gecko dataset with different tilts and apertures.

Based on the incremental convolution (*), we calculate a modulation factor λ_i for each sample on the slice i which determines the visibility of the current slice:

$$\lambda_i = \frac{1}{1 + G_\epsilon * O_{i-1}} \quad (13)$$

In addition to the opacity buffer O_i , we use a color buffer C_i for each slice. The opacity buffer for the next slice combines the opacity buffer of the previous slice with the opacity of the current slice α_i :

$$O_i = G_\epsilon * O_{i-1} + \alpha_i \quad (14)$$

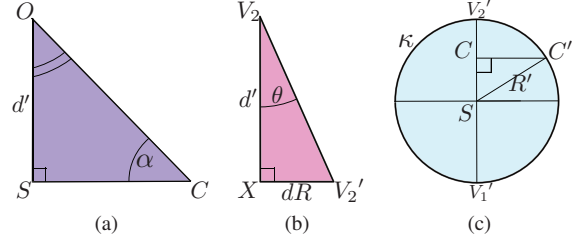


Figure 6: We introduce new literals for selected primitives: (a) the triangle $\triangle SCO$, (b) the triangle $\triangle XV_2'V_2$ and (c) the circle κ . These primitives are defined in Figures 4 and 5 by the same color encoding. Note, that $\|CC'\| = B$ which is the minor axis of the ellipse ϵ .

The color contribution c_i is multiplied by λ_i . The color c_i and opacity α_i propagate to C_{i+1} using traditional alpha-blending with the over operator. Our method requires no pre-computation and performs at interactive frame-rates. Due to incremental blurring of the opacity buffer, shadows cast by highly occlusive regions fade-out smoothly with distance. Compared to the state-of-the-art model, we thereby add a movable light source with negligible performance penalty.

4. Implementation Details

Our new model was implemented as a plugin to VolumeShop [BG05] using C++ and OpenGL/GLSL. Using the ARB_draw_buffers OpenGL extension, two render targets are written for each slice: the intermediate rendered image and the occlusion buffer. The elliptical blurring kernel is stored in an additional texture which is updated whenever the light source parameters change. For all examples in the paper, we use a texture size of 128×128 . When the lighting parameters change, we recompute the footprint. The major axis of the ellipse is aligned with the projection of the light vector to the viewing plane by multiplying GL_TEXTURE matrix stack by a rotation matrix. In case the ellipse grows or moves out of the texture, we apply translation and scaling to

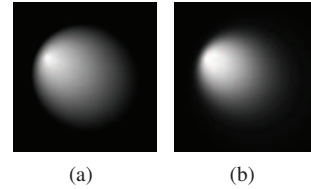


Figure 7: Elliptical kernels used for incremental blurring of the opacity buffer: with linear fall-off (a) and Gaussian fall-off of the weighting function (b).

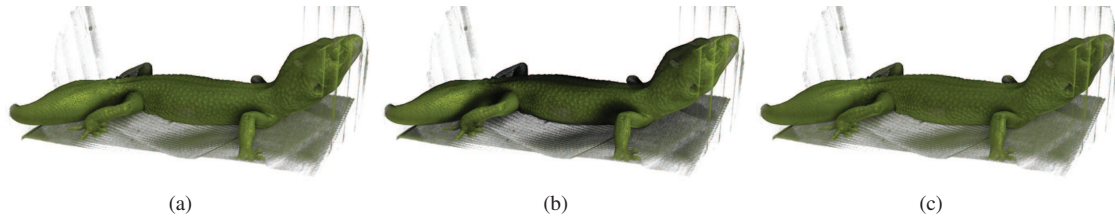


Figure 8: Visualizations of the gecko CT dataset with different setup of aperture θ and tilt angle α : $\theta = 10^\circ$ and $\alpha = 37^\circ$ (a), $\theta = 40^\circ$ and $\alpha = 37^\circ$ (b), and $\theta = 40^\circ$ and $\alpha = 5^\circ$ (c). Light is coming from top-right for all images.

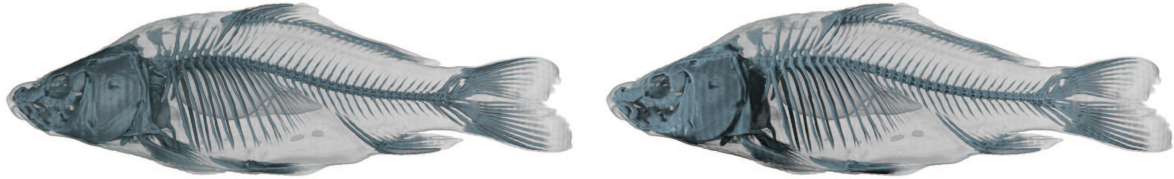


Figure 9: Visualizations of the carp CT dataset using the directional occlusion shading model with a headlamp illumination setup (a) and using illumination setup conventional to medical illustrations (b).

fit it into the bounding box of the texture. During rendering, the inverse transformation is applied to access the kernel at correct positions. However, massive downscaling of the coordinate system may lead to a loss of precision. Users can interactively adjust the tilt, the aperture, and the XY-rotation of the light source. This gives the user full control to set the light source arbitrarily in the hemisphere defined by the view-vector. The parameters aperture, tilt, and rotation are set by sliders in the user-interface.

5. Results and Discussion

In this section, we provide case-studies and comparisons to other volume rendering approaches and analyze the performance of our new method.

5.1. Case Studies

Medical illustrators generally place the light source in a top-left corner to improve depth perception. Figure 9 depicts the carp CT dataset visualized under different illumination conditions. While in Figure 9(a), the image lacks depth, Figure 9(b) emphasizes the detailed structure of the skeleton through shadows. Similarly, Figure 10 shows cases where illumination leads to better perception of structures. In Figure 10(a), the hand seems to directly contact the body. In reality, there is a small gap which is visible in Figure 10(b). Similarly for Figures 10(c) and 10(d): in Figure 10(d), the eye sockets of the skull appear deeper than in Figure 10(c). We consulted a certified medical illustrator with over 25

years of professional experience who affirmed that the visualizations generated using this kind of illumination yield stronger perception cues. We presented her visualizations using different lighting settings. Her task was to choose which lighting conditions suit medical illustrations the best. She consistently preferred image such as those depicted in Figures 10(b) and 10(d). The illustrator further confirmed that interactive fine-tuning of the exact light placement is necessary in many cases, in order to avoid excessive darkening of focus objects. In volume data, regions with high gradient magnitude correspond to surface-like structures. Using the gradient magnitude to add an additional local specular component to these objects can further improve perception. Figure 11 presents a computer tomography of a human foot generated with different illumination models and varying light source positions: Figures 11(a) and 11(b) use the multidirectional OS model enhanced by specular highlights, and Figures 11(c) and 11(d) use the pure multidirectional OS model.

To gain a coarse impression on the impact of our technique on non-professionals, we also conducted a small user study on a group of 42 participants with different backgrounds. We presented them two series of result images: the human hand and the human thorax which are shown in Figures 1 and 12. Their task was to choose an image which in their opinion yields the strongest depth cues. From the series of different renderings of the hand, 39 participants (92.86%) favored the top-left illumination in Figure 1(d), 2 participants (4.76%) preferred the raycasting in Figure 1(a) and 1 participant (2.38%) preferred the head-lamp illumination in Figure 1(c). A majority of 41 (97.62%) participants also pre-

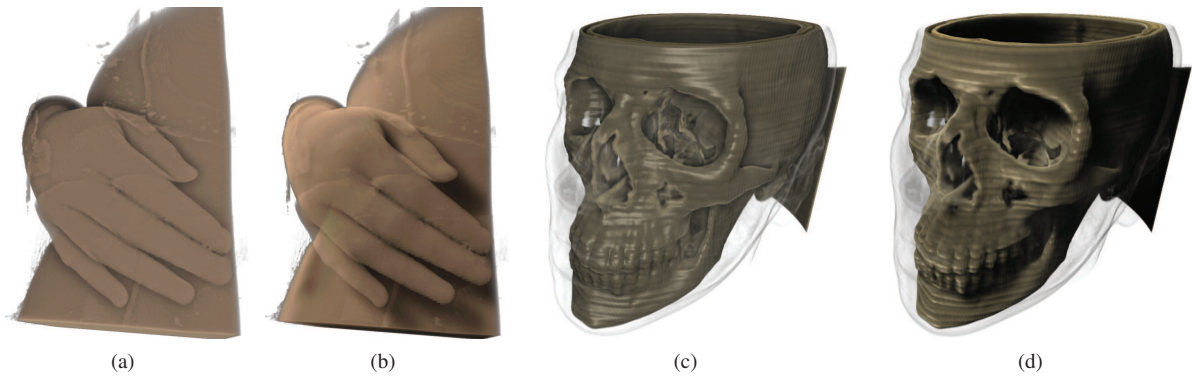


Figure 10: Visualizations of computer tomography data using the directional occlusion shading model with a headlamp illumination setup (a) and (c) using the illumination setup conventional to medical illustrations (b) and (d).

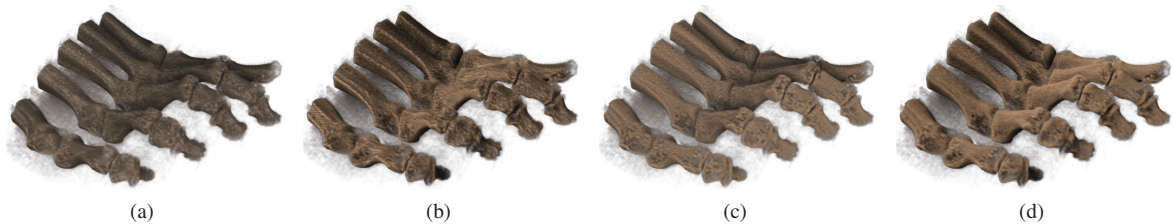


Figure 11: Visualizations of a human foot acquired by computer tomography using the directional occlusion shading model: using the Phong illumination model with the headlamp illumination setup (a) and with the top-left lighting (b). Visualizations (c) and (d) use the diffuse illumination model with the headlamp and the top-left light-source setup respectively.

ferred the top-left illumination of the thorax in Figure 12(d) and only one participant (2.38%) selected the raycasted image in Figure 12(a).

Local surface-based illumination of volume data employs the gradient to substitute for the surface normal. However, gradients estimation frequently performs poor in the presence of noise which can lead to distracting artifacts. Thus, for modalities such as ultrasound, unshaded volume rendering is commonly employed. This makes the structures in the data difficult to interpret even for experienced medical professionals. The directional occlusion shading model as a gradient-free shading method can be used to improve perception. Interactive light source modification enables the user to inspect and understand the structures better. Figure 13 shows different visualizations of 3D cardiac ultrasound: 2D slices and 3D volume renderings. The clipping plane reveals the inside of the heart chambers. During examination, physicians see the ultrasound visualizations on their workstations as in Figure 13(a). We used a transfer function which shows the heart in a similar fashion. Figure 13(b) shows that gradient-based shading is not well-suited for ultrasound data. Multidirectional occlusion shading, on the other hand, reveals the structure, and interaction with light

source enables the user to better perceive the depth of the cavities.

We described a shading model which does not require precomputation and storage of additional data, unlike deep shadow maps [HKS06] or light volumes [RDR10], and which allows arbitrary light position within the hemisphere defined by the view vector. Half-angle slicing, introduced in the work of Kniss et al. [KKH02], generates shadows by using a slicing direction halfway between view and light direction. However, choosing a slicing direction which is non-parallel to the viewing directions leads to visible artifacts, especially when the light source tilt angle surpasses 60° . Figure 14 clearly demonstrates a situation when such artifacts are visible when half-angle slicing is used. In the original half-angle slicing approach, the order of the slices is reverted if the light source is located in the hemisphere opposite to the viewer. Reverting the traversal of the slice-stack is a possible extension of our approach which would not limit the light vector to the hemisphere defined by the view vector.

5.2. Performance Analysis

We tested the described method on a workstation equipped with an NVIDIA GeForce 295 GTX GPU with 1.7GB graph-

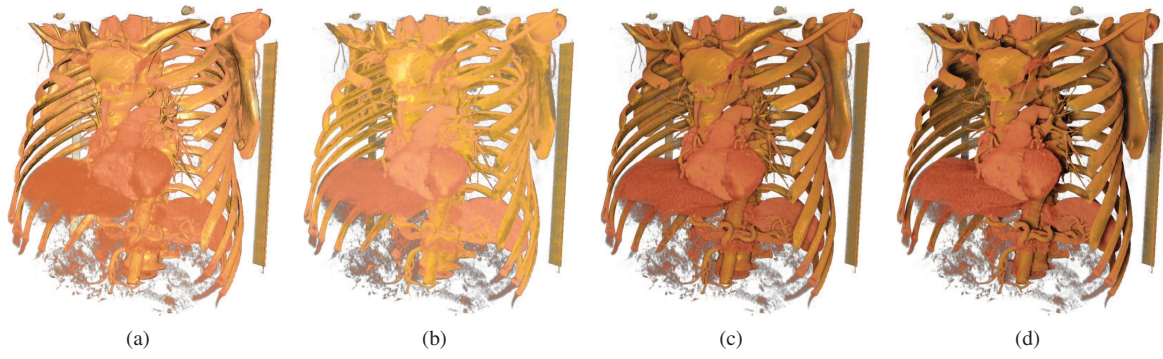


Figure 12: Visualizations of a human thorax we used for user study: using raycasting (a), sliced-based volume rendering (b), both using the Phong illumination followed by the directional occlusion shading model with the headlamp illumination setup (c) and illuminated from top-left (d).

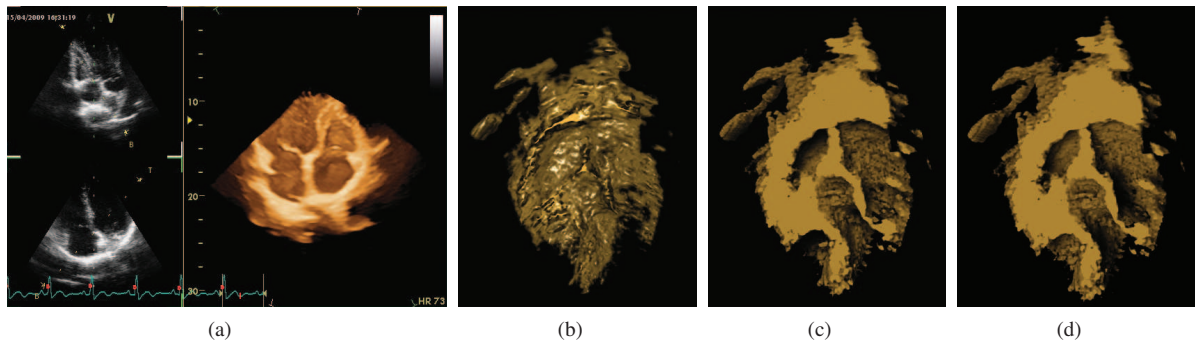


Figure 13: Visualizations of 3D ultrasound of cardiac data: user interface of a 3D cardiac ultrasound workstation (a), a clipped 3D cardiac ultrasound visualization using direct volume rendering and Phong illumination model, rendered with a raycaster (b), clipped 3D cardiac ultrasound visualization using the multidirectional occlusion shading model with light coming from top-left (c) and bottom-left (d).



Figure 14: Visualizations of the bonsai dataset: slice-based volume rendering using a view-aligned slice stack (a) and using a half-angle-aligned slice stack (b).

ics memory, an Intel®Core i7 CPU with 3.07GHz and 12GB of RAM. We measured the performance of our implementation using the gecko-dataset of resolution $512 \times 512 \times 88$ voxels, 0.5 voxels sampling distance and viewport-resolution 768×407 pixels. We achieved interactive frame-rates of 19Hz with using the MS-DOS and 18Hz with mul-

tidirectional OS using a 37° angle of aperture while interacting with the viewing parameters. During interaction with the light source, which required update of the kernel, we achieved 14Hz frame-rates. For comparison, using the same framework, a simple slice-based renderer with no shadowing and Phong illumination achieved 25Hz and a high-quality raycaster with Phong illumination and no shadowing achieved 21Hz. We performed the same test with the foot dataset of resolution $256 \times 256 \times 256$ voxels, 0.5 voxels sampling distance and viewport-resolution 531×311 pixels. We achieved 15Hz while using the original MS-DOS approach, 14Hz using our new method, and 12Hz during light-source interaction. In this case, a simple slice-based renderer performed at 25Hz and a raycaster at 22Hz. These tests prove that the interactive light-source placement is a valuable extension of the original approach traded for a negligible performance penalty.

6. Conclusions

In this paper, we presented a shading model for direct volume rendering, which enables the interactive generation of high-quality soft shadow effects without the need for pre-computation. Our method extends a previous technique to enable interactive placement of the light source. Using elliptic instead of circular footprints, we achieve almost the same performance while greatly improving the flexibility of the method. Additionally, we discussed several applications of such a shading model and consulted a professional illustrator to confirm the importance of freely modifying the light direction.

7. Acknowledgments

This work was carried out within the IllustraSound research project (# 193170), which is funded by the VERDIKT program of the Norwegian Research Council with support of the MedViz network in Bergen, Norway. The authors wish to thank the certified medical illustrator Kari Toverud for consulting, numerous respondents for their feedback and anonymous reviewers for their comments.

References

- [BBC83] BERBAUM K., BEVER T., CHUNG C. S.: Light source position in the perception of object shape. *Perception* 12, 5 (1983), 411–416.
- [BG05] BRUCKNER S., GRÖLLER M. E.: VolumeShop: An interactive system for direct volume illustration. In *Proceedings of IEEE Visualization* (2005), pp. 671–678.
- [BG07] BRUCKNER S., GRÖLLER M. E.: Enhancing depth-perception with flexible volumetric halos. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1344–1351.
- [BLK00] BRAJE W., LEGGE G., KERSTEN D.: Invariant recognition of natural objects in the presence of shadows. *Perception* 29, 4 (2000), 383–398.
- [BR98] BEHRENS U., RATERING R.: Adding shadows to a texture-based volume renderer. In *Proceedings of IEEE Symposium on Volume Visualization 1998* (1998), pp. 39–46.
- [Bun05] BUNNEL M.: *GPU-Gems*, vol. 2. Addison Wesley, 2005, ch. Dynamic Ambient Occlusion and Indirect Lighting.
- [DEP05] DESGRANGES P., ENGEL K., PALADINI G.: Gradient-free shading: A new method for realistic interactive volume rendering. In *Proceedings of Vision, Modeling, and Visualization 2005* (2005), pp. 209–216.
- [DSDD07] DACHSBACHER C., STAMMINGER M., DRETTAKIS G., DURAND F.: Implicit visibility and antiradiance for interactive global illumination. *ACM Transactions on Graphics* 26, 3 (2007), 61.1–61.10.
- [HKS06] HADWIGER M., KRATZ A., SIGG C., BÜHLER K.: GPU-accelerated deep shadow maps for direct volume rendering. In *Proceedings of SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware* (2006), pp. 49–52.
- [HLY08] HERNELL F., LJUNG P., YNNERMAN A.: Interactive global light propagation in direct volume rendering using local piecewise integration. In *Proceedings of IEEE/EG International Symposium on Volume and Point-Based Graphics* (2008), pp. 105–112.
- [HLY09] HERNELL F., LJUNG P., YNNERMAN A.: Local ambient occlusion in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics* 9, 2 (2009).
- [HLY07] HERNELL F., YNNERMAN A., LJUNG P.: Efficient ambient and emissive tissue illumination using local occlusion in multiresolution volume rendering. In *Proceedings of Volume Graphics 2007* (2007), pp. 1–8.
- [KKH02] KNISS J., KINDLMANN G., HANSEN C.: Multidimensional transfer functions for interactive volume rendering. *IEEE Transactions on Visualization and Computer Graphics* 8, 3 (2002), 270–285.
- [KMK94] KERSTEN D., MAMASSIAN P., KNILL D.: *Moving cast shadows and the perception of relative depth*. Tech. Rep. 6, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany, 1994.
- [Kne07] KNECHT M.: *State of the art report on ambient occlusion*. Tech. rep., Technische Universität Wien, Vienna, Austria, 2007.
- [KPH*03] KNISS J., PREMOZE S., HANSEN C., SHIRLEY P., MCPHERSON A.: A model for volume lighting and modeling. *IEEE Transactions on Visualization and Computer Graphics* 9, 2 (2003), 150–162.
- [Lev87] LEVOY M.: Display of surfaces from volume data. *IEEE Computer Graphics and Applications* 8 (1987), 29–37.
- [Max95] MAX N.: Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics* 1, 2 (1995), 99–108.
- [MS09] MÉNDEZ-FELIU À., SBERT M.: From obscurities to ambient occlusion: A survey. *"The Visual Computer"* 25, 2 (2009), 181–196.
- [RDR10] ROPINSKI T., DÖRING C., REZK-SALAMA C.: Interactive volumetric lighting simulating scattering and shadowing. In *Proceedings of IEEE Pacific Visualization* (2010), pp. 169–176.
- [RMSD*08] ROPINSKI T., MEYER-SPRADOW J., DIEPENBROCK S., MENSMAJN J., HINRICHS K. H.: Interactive volume rendering with dynamic ambient occlusion and color bleeding. *Computer Graphics Forum* 27, 2 (2008), 567–576.
- [RS07] REZK-SALAMA C.: GPU-based Monte-Carlo volume raycasting. In *Proceedings of Pacific Graphics 2007* (2007), pp. 411–414.
- [SA07] SHANMUGAM P., ARIKAN O.: Hardware accelerated ambient occlusion techniques on GPUs. In *Proceedings of Symposium on Interactive 3D Graphics and Games* (2007), pp. 73–80.
- [SPH*09] SCHOTT M., PEGORARO V., HANSEN C., BOULANGER K., STRATTON J., BOUATOUCH K.: A directional occlusion shading model for interactive direct volume rendering. *Computer Graphics Forum* 28, 3 (June 2009), 855–862.
- [Ste03] STEWART A. J.: Vicinity shading for enhanced perception of volumetric data. In *Proceedings of IEEE Visualization 2003* (2003), pp. 355–362.
- [YKZ91] YAGEL R., KAUFMAN A., ZHANG Q.: Realistic volume imaging. In *Proceedings of IEEE Visualization 1991* (1991), pp. 226–231.
- [ZIK98] ZHUKOV S., IONES A., KRONIN G.: An ambient light illumination model. In *Rendering Techniques* (1998), pp. 45–56.

2.2 Illustrative Membrane Clipping

Åsmund Birkeland, Stefan Bruckner, Andrea Brambilla, and Ivan Viola

Computer Graphics Forum, Volume 31, Issue 3pt1, pages 905–914, 2012

DOI: 10.1111/j.1467-8659.2012.03083.x

The definitive version is available at <http://diglib.eg.org/>
and <http://onlinelibrary.wiley.com/>.

Illustrative Membrane Clipping

Å. Birkeland¹, S. Bruckner², A. Brambilla¹ and I. Viola^{1,3}

¹University of Bergen, Norway

²Vienna University of Technology, Austria

³Christian Michelsen Research, Norway

Abstract

Clipping is a fast, common technique for resolving occlusions. It only requires simple interaction, is easily understandable, and thus has been very popular for volume exploration. However, a drawback of clipping is that the technique indiscriminately cuts through features. Illustrators, for example, consider the structures in the vicinity of the cut when visualizing complex spatial data and make sure that smaller structures near the clipping plane are kept in the image and not cut into fragments. In this paper we present a new technique, which combines the simple clipping interaction with automated selective feature preservation using an elastic membrane. In order to prevent cutting objects near the clipping plane, the deformable membrane uses underlying data properties to adjust itself to salient structures. To achieve this behaviour, we translate data attributes into a potential field which acts on the membrane, thus moving the problem of deformation into the soft-body dynamics domain. This allows us to exploit existing GPU-based physics libraries which achieve interactive frame rates. For manual adjustment, the user can insert additional potential fields, as well as pinning the membrane to interesting areas. We demonstrate that our method can act as a flexible and non-invasive replacement of traditional clipping planes.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Volume Rendering—Manipulation and Deformation

1. Introduction

Clipping is an essential mechanism in visualization which resolves spatial occlusion in three- or more dimensions. This basic idea is used in visual depiction in many different variants, such as clipping into half-spaces, near clipping planes, section views, and cutaway-views which can all be considered variants of the concept.

In 3D data visualization, clipping becomes very powerful when making it adjustable by the user or adapting it to data properties. Clipping provides a simple mechanism for revealing otherwise invisible structures. Similarly, slicing uses the same interaction scheme, but shows only a single slice without any context outside of the intersecting plane.

The idea of removing occluding objects has been used in illustration craft for centuries. However, at closer inspection traditional illustrations rarely employ strict planar clipping but actually use hybrid representations where particular structures extend in front of and behind a hypothetical

selected plane. Similarly, elongated structures do not appear or disappear from the rendering, as they would when cut by a plane. Instead, they are aligned with the cut plane.

One example, are dense cell illustrations where structures are clearly discernible and emerge out of the hypothetical clipping plane. It is then possible to follow elongated structures even if they spatially overlap. Figure 1 shows an illustration of a slice through a *Mycoplasma mycoides* cell. As one can see in this illustration, the cell appears to be cut open by a plane revealing the inner structures, yet the structures at the position of the cutting plane are not sliced.

This is practically impossible to achieve with dense 3D visualizations. Objects will likely be intersected by the clipping plane so that only partial structures are depicted. The remaining portion itself may not have a clear meaning to the viewer and may hinder spatial comprehension. The research question arises of how to formalize the illustrative clipping concept in an explorative 3D data visualization scenario with limited semantic information at hand. One way of interpret-

ing illustrative clipping is that the clipping geometry is not planar, but adapts to the structure of the object. The resulting clipping geometry will be a deformed 2D manifold instead of a plane. This manifold preserves the consistency in the structural depiction, provides subtle 3D shape cues and allows the following of elongated structures over the illustration. The resulting clipping membrane reveals otherwise occluded structures but at the same time communicates additional 3D information.

In our approach, the user positions the membrane in the scene and the scene elements act on the clipping geometry to cause the deformation. The membrane is utilized in order to reveal more of its structural properties, and to remove irrelevant information. To avoid cutting through an object when the focus is on its outer shape rather than its inner structures, the object repels the clipping membrane. Other data properties, such as object interfaces, can act attractively, so that the clipping membrane tends to be aligned with these interfaces. For this purpose, our clipping membrane concept is translated to a physics optimization problem. The scene elements generate a potential field where each data element's force is specified by a *potential field transfer function* (PFTF). The resulting potential field is an integral of (signed) forces which act on a deformable cloth object that constitutes the membrane. The membrane is not explicitly visible, but serves as a selection tool for direct volume rendering (DVR).

This idea has a wide range of potential applications. One scenario is the enhancement of traditional 2D slicing with selective depiction of 3D structures, but it also applies to adaptive clipping during mechanic movements of imaged structures in time-dependent data. To achieve interactive performance, the implementation of our clipping membrane concept makes use of NVIDIA's PhysX technology, which executes a physically-based cloth simulation on the GPU.

2. Related Work

3D region extraction or clipping are common tools in medical workstations, for instance in ultrasound scanners with 3D acquisition capability [Kre99]. Manually defining the clipping geometry has been very popular for creating expressive visualizations [WEE02, NWKY05, FWH09]. Konrad-Verse et al. showed how manual sketching can define a deformable cutting plane, aiding in virtual resection [KVLPO4]. Deformed planes are also used in slice-based blood-vessel visualization, reforming a plane to the center of the vessel [KFW*02]. However, manually defining complex geometries can be cumbersome. Automatic definition of 3D clipping regions is typically based on semantic information [VKG04].

Selective clipping adjustment has also been used in several visualization techniques. For preoperative planning, Beyer et al. demonstrated the use of clipping for combining

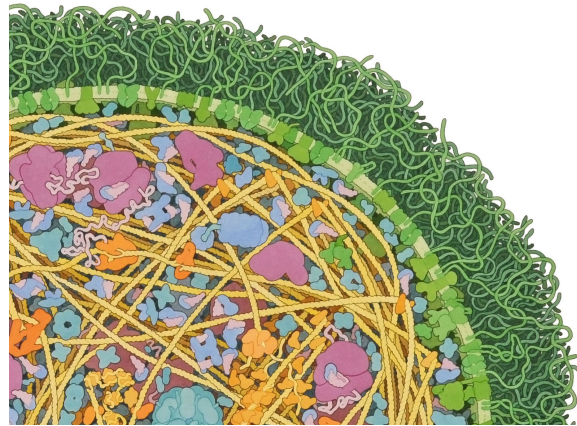


Figure 1: Illustration using clipping to reveal the macro-molecular structures inside a *Mycoplasma mycoides* cell, by David S. Goodsell [goo12], the Scripps Research Institute.

the information of multiple imaging modalities [BHWB07]. Similarly, Burns et al. have developed a technique for combining 2D ultrasound and 3D CT in a fused visualization [BHW*07]. In their approach, the clipping geometry was defined by a 2D ultrasound plane. For additional context, certain features defined by segmentation were not clipped. Inspired by enhanced slice rendering [TMS*06] where contextual information of the surrounding areas is merged with the slice, we propose a method for using a deformed geometry to create expressive slice rendering without any required segmentation. Volume splitting [ISC07] can also be seen as an advanced form of clipping, but rather than removing parts of the volume, it translates the clipped parts in order to ensure direct line of sight. An example of this are exploded views [BG06], where the different parts of the volume are repelled from each other, as well as repelled from the line of sight.

Schultz et al. used deformable clipping geometry to simulate a familiar resection method for nerve tracts in the brain, called Klingler resection [SSA*08]. The technique known as Virtual Klingler uses the data intensities to aid in the deformation of the clipping geometry. Schultz's method was the first one to use forces to deform a clipping geometry. We see this method as a specific application of the general concept presented in this paper. While the Virtual Klingler method was highly focused on producing visualizations similar to Klingler resection, the technique had some limitations. The deformation of the plane was defined by creating a resistance field from MRI datasets. To achieve a deformation, the plane needed to be manually moved through the volume. We propose to use active forces, derived from the data, to deform an elastic geometry employed in volume visualization. In addition, we provide a controlling mechanism connecting data attributes to the strength of the forces and propose interaction schemes as well as visual mapping strategies.

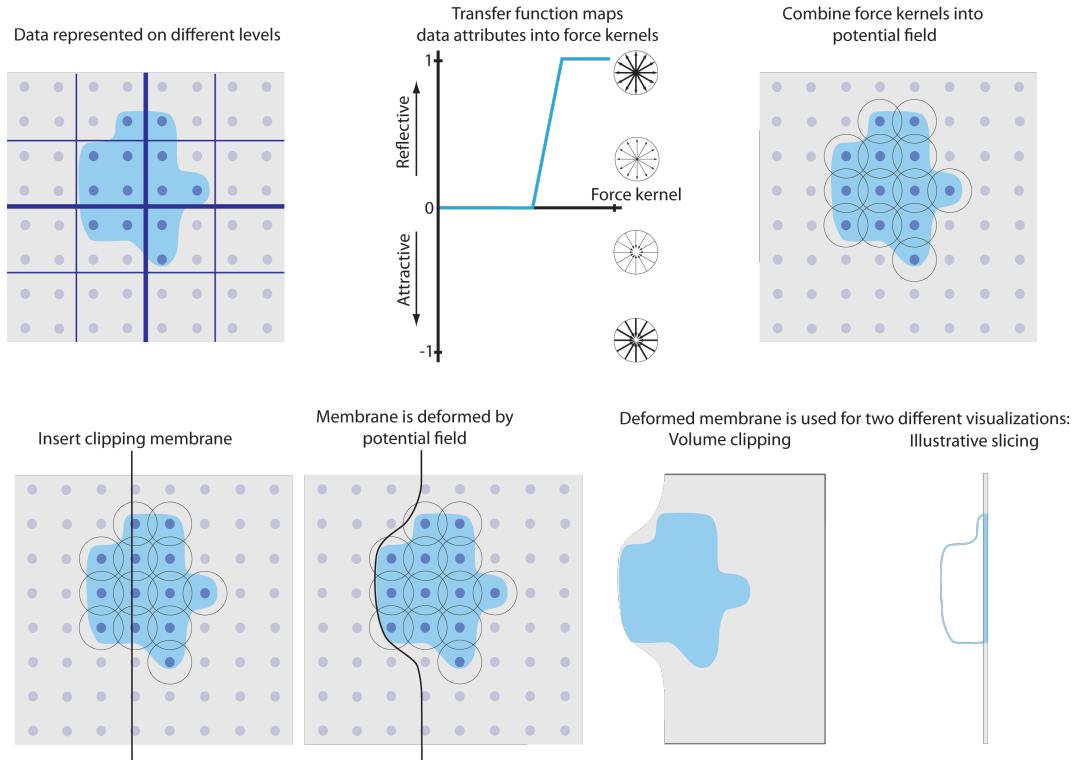


Figure 2: Elementary building blocks of the illustrative clipping pipeline and the two use-case scenarios. Data samples from a specific level are mapped to forces using a *potential field transfer function*. The forces constitutes a potential field, which interact with a flexible membrane. The deformed membrane is then applied in volume visualization.

Physically-based deformation of three-dimensional models is a very active research area, especially in the context of computer graphics. A detailed overview of the most recent approaches in this field was presented by Nealen et al. [NMK*06]. A notable amount of work has been dedicated to the physical simulation of cloths [HB00, MTCK*04]. In our work, we exploit the features of the NVIDIA PhysX library to simulate an elastic piece of cloth which acts as a clipping plane. Details about how the library handles this simulation can be found in Section 3.1 or in the paper by Müller et al. [MHHR07].

3. Methodology

The core idea of this paper, is to translate the illustrative clipping concept into an interaction between soft-body dynamics actors, as depicted in Figure 2. Firstly, the data attributes are translated into physical elements. This is done by creating an attractive or repelling force for each data element, called a *force kernel*. The combined forces from all the data elements construct a *potential field* depending on the data attributes. In addition, the user needs a controlling mechanism for linking data attributes to the strength of forces exerted by an element. We adapted the familiar transfer function concept,

into a *potential field transfer function* (PFTF). Data attributes are then mapped by the PFTF to the strength of the force kernels.

Secondly, we need a geometry which can interact with the forces in the potential field. We therefore create a mesh which acts as an elastic membrane and then insert it into the potential field. By applying a soft-body physics simulation, each vertex in the membrane interacts with the nearby force kernels and conforms to the potential field created from the data attributes. As we aim to use the deformed geometry as a means to create expressive volume visualizations, we propose two different methods to exploit the deformable geometry in already existing rendering schemes. In volume clipping, we exchange the regular flat clipping plane with our deformed geometry. As a second utility example, we suggest the use of the deformed geometry for illustrative slice rendering, enhancing it with the local context defined by the deformed membrane.

3.1. Deformable Membrane

The clipping membrane is represented as a constrained elastic piece of cloth whose shape and position are automati-

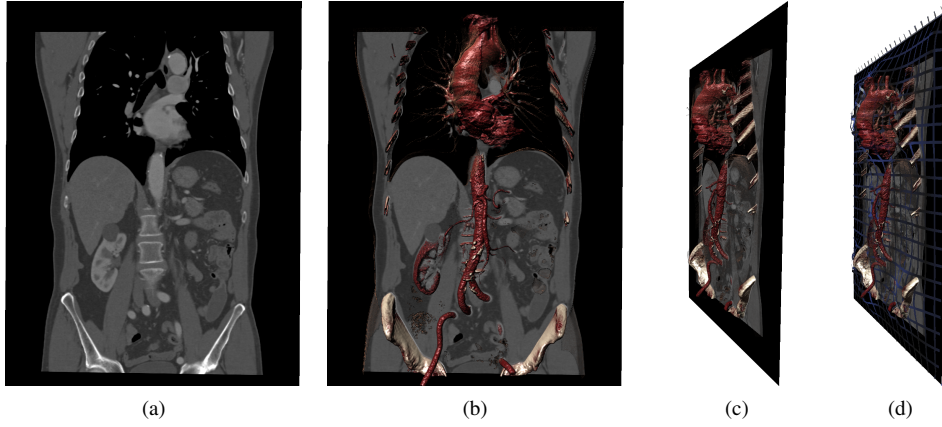


Figure 3: Rendering of a medical CT dataset. Compared to the regular slice view (3a), adding local DVR provides interesting structures out of the plane (3b). Figure 3c shows a tilted view of the same plane and the grid in Figure 3d shows the deformed membrane.

cally determined according to a set of physical properties. The physical simulation is based on *Position Based Dynamics* (PBD) [MHHR07], a physics-based approach to simulate soft bodies. Typical physical simulation algorithms are based on Newton's second law of motion, where every particle is associated with an ordinary differential equation (ODE) which includes the forces acting on the particle, the particle's velocity and acceleration, and possibly additional constraints. The system of ODEs is usually solved using an integration algorithm that takes into account all these variables. In the PBD approach, the positions of the particles are directly manipulated: in this way overshooting problems (caused by strong accelerations or high velocities) are avoided, and collisions and penetrations can be easily managed.

In PBD, the dynamical system is defined by a set of vertices and a set of constraints. Similarly to the common mass-spring models, every vertex $i \in \{1, \dots, N\}$ has a mass m_i , a position \mathbf{x}_i and a velocity \mathbf{v}_i . A constraint $j \in \{1, \dots, M\}$ instead is described by either an equality $C_j(x_{i_1}, \dots, x_{i_{n_j}}) = 0$ or an inequality $C_j(x_{i_1}, \dots, x_{i_{n_j}}) \geq 0$, where C_j is a real valued function defined over the positions of a subset of the particles. A stiffness parameter $k_j \in [0, 1]$ determines the *strength* of the constraint. Given a time step size Δt , the vertices are initially advected according to the explicit Euler's integration scheme [HNW93]:

$$\begin{cases} \tilde{\mathbf{v}}_i &= \mathbf{v}_i + \Delta t \frac{\mathbf{f}_{ext}(\mathbf{x}_i)}{m_i} \\ \tilde{\mathbf{x}}_i &= \mathbf{x}_i + \Delta t \tilde{\mathbf{v}}_i \end{cases} \quad (1)$$

where $\mathbf{f}_{ext}(\mathbf{x}_i)$ is the force the potential field exerts on particle i . At this point collisions are detected and, for each of them, an additional constraint is generated. Finally, the algorithm iterates through all the constraints and repeatedly

modifies the temporary particle positions $\tilde{\mathbf{x}}_i$ trying to satisfy all the equalities and inequalities (taking into account their stiffness k_j). For more detailed information about the general approach, we refer to the article of Müller et al. [MHHR07].

In the case of cloth simulation, two types of constraints are used: the *stretching* and *bending*. The stretching constraint is defined over every edge of the cloth mesh. It aims at preserving the original length of the edge, therefore the associated stiffness parameter can be used to control how rigid the cloth is. The bending constraint instead preserves the angle between each pair of adjacent triangles, so its stiffness parameter determines the cloth's resistance to folding.

To maintain the simple clipping interaction the membrane must retain a certain shape. Without any constraints, the membrane would collapse. In our technique, the membrane is attached to a rigid frame. The rigid rectangular frame ensures that the membrane retains a topologically planar shape.

3.2. Potential field

The potential field is essentially a function $\mathbf{f}: \{1, \dots, N\} \rightarrow \mathbb{R}^3$ which associates a force vector to each of the N particles in the dynamical system. In particular, in the NVIDIA PhysX framework, the potential field is given by the sum of forces \mathbf{f}_k exerted by a set of *kernels* $k \in \{1, \dots, K\}$ on the particles:

$$\mathbf{f}(i) = \sum_{k=1}^K \mathbf{f}_k(i) \quad (2)$$

The framework lets the user freely specify how the kernel force functions are computed. Specifically, in order to support datasets based on non-cartesian grids, we decided

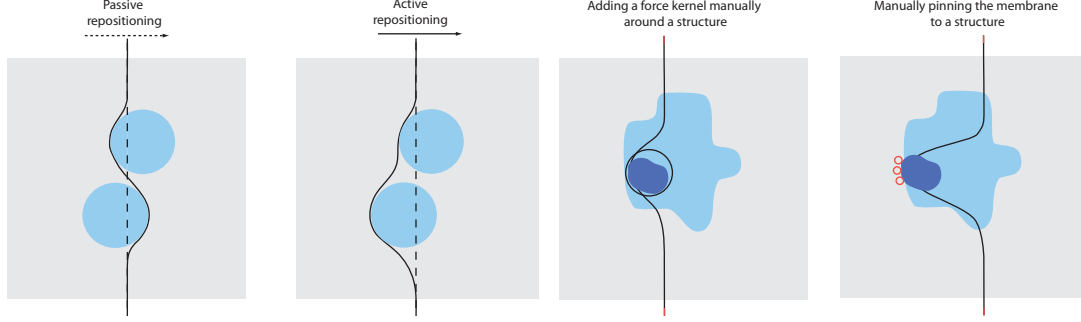


Figure 4: Two different methods for repositioning the membrane: Passive repositioning creates a new membrane at the position of the rigid frame. Active repositioning moves the frame and pulls the membrane through the potential field. We provide a means for interacting locally with the potential field by creating a force kernel with a desired radius to encapsulate interesting structures. In addition, the user can pin the membrane at the current position to prevent the membrane from moving through interesting features.

to treat each of the three components of \mathbf{f}_k separately. The force a kernel k , with position \mathbf{x}_k , exerts on a particle i , with position \mathbf{x}_i , is given by:

$$f_k^j(i) = c(1 - \frac{|\mathbf{x}_k - \mathbf{x}_i|}{a_j}) \quad (3)$$

with $j \in \{0, 1, 2\}$. The parameter c is determined using the PFTF, while the parameter a_j defines the area of effect of the kernel and is proportional to the voxel size in the j^{th} direction.

At this point, building the potential field is straightforward: for every voxel the PFTF is evaluated, then a suitable kernel is generated and placed at the voxel's center. However, even with moderate size datasets, the resulting field would be so complex that the physical simulation would not run at interactive rates. To make the system more scalable, we adopt an octree structure to control kernels' placement: the desired octree depth is specified by the user, then, for every super-voxel at that depth, a kernel is generated. The kernel is placed at the center of the super-voxel and its strength is obtained averaging the PFTF values of all the included voxels. Eventually, the force exerted by the potential field on a certain particle is taken into account during the simulation as the f_{ext} term of Equation 1. The equilibrium state of the membrane is then determined through the PBD algorithm.

While the potential field push the membrane away from the original slice position, the only force back towards the slice comes from stretching the nodes. This results in the membrane not following the the cavities of the potential field. As a means to force the membrane back towards its original position, we add a global force towards the slice, resulting in a tighter fit around the structures.

3.3. Interaction

The main interaction with the potential field is done using a widget for creating a transfer function. Adjusting the transfer function to data with very similar data values can be very challenging. To simplify interaction, we allow the user to sketch directly on the original slice or clipping plane. From the sketch, we define a neighbourhood and create a histogram of the included samples which we use to adjust the potential field transfer function.

Positioning the membrane is essentially as simple as moving an ordinary clipping plane or slice through the volume. We propose two different schemes for positioning the membrane, passive insertion and active movement, as illustrated in Figure 4. During passive insertion, the membrane is not affected by the force field as it moves. In essence, a membrane is created inside the potential field during every reposition, and it is then deformed by the force kernels it intersects. In active mode, the membrane is moved across the potential field, and it is deformed in real-time during movement. As the membrane moves across the volume, it *collides* with the potential field. While the frame is moved further across the volume, the membrane moves through the kernels from one equilibrium state to another.

For certain structures, very similar data values makes it difficult defining a proper PFTF. A more direct interaction with the membrane can be useful to keep interesting features in view. We have created two methods for manually interacting with the physical model. One method, is where the user indirectly deforms the membrane by inserting additional repulsive or attractive force kernels at the location of the original slice. Then, using a simple click-and-drag motion, the user can add a kernel with the desired radius into the existing potential field at the depth specified by the slice position.

In addition, the user might desire to keep certain struc-

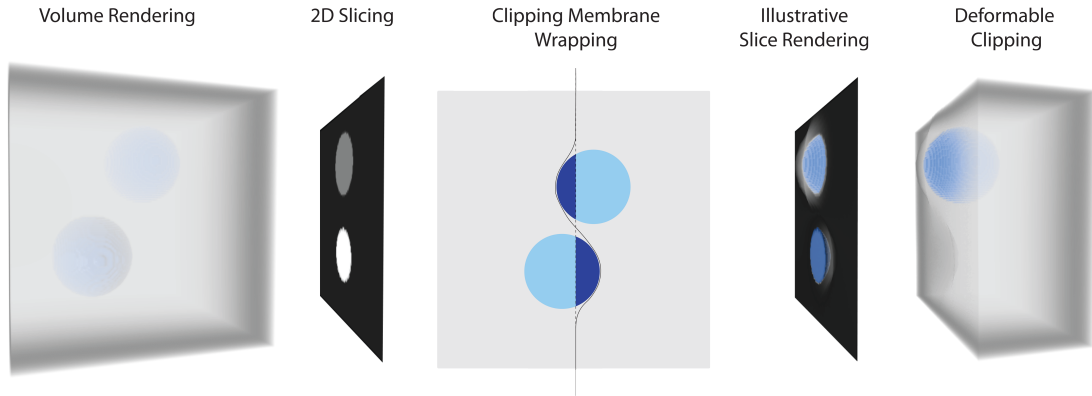


Figure 5: Traditional explorative tools compared to illustrative clipping applied to a synthetic dataset.

tures in view while exploring the rest of the volume. In this case, the membrane can be *pinned* to structures by simply clicking on the membrane. An immovable anchor is then attached to the membrane, keeping the attached elements from moving with the rest of the membrane. This allows the user to manually *fine tune* the clipping. Figure 4 shows examples of the two interaction schemes. To better illustrate the effect, we removed all other forces from scene.

3.4. Visual Mapping

We propose to apply the deformed clipping geometry in two different ways: membrane clipping in DVR and illustrative slicing with local DVR. Figure 5 illustrates the different rendering methods. In both cases, we use a bounding geometry, for calculating the entry and exit points for ray casting.

Regular clipping in ray casting can be achieved using a simple dot product test. With a deformed geometry, it becomes more challenging. Instead of testing which elements are in front or behind the plane, we start the ray casting at the surface of the cut. We then replace one face of the volume's bounding box with the clipping membrane. The new bounding geometry can be used to calculate new entry and exit points from the position of the membrane.

While clipping is a very common technique for volume exploration, domain experts, especially in radiology, tend to use slice rendering as their main visualization tool. Slice rendering provides highly detailed views, with no occlusion. The slice contains little or no context outside of the plane and a challenge rests on the user to comprehend the outer planar spatial extent of structures. Slab rendering can overcome this problem to some extent, by choosing a subset of the volume to be rendered. A slab is a subset of the volume defined by two parallel clipping planes. Similarly, with regular clipping, uninteresting objects may clutter the image and interesting features can be clipped.

To add 3D context to the slice of only those structures

the user finds interesting, we propose using the elastic membrane for illustrative slicing, where DVR is only added in between the deformed membrane and the original slice. In this manner, we calculate the entry and exit points for ray casting from the proxy-geometry created by combining the slice and the membrane. The interesting structures are typically in the vicinity of the plane. We prevent clutter from distant structures by modulating the opacity of a certain sample, by the distance to the plane, as depicted in Figure 6. Blending between the slice and the local DVR, it is important to be able to differentiate between structures in front or behind the plane. We have created a technique for user-selected blending and emphasis. Using two sliders, the user selects the opacity and the emphasis between structures behind and in front of the plane. One slider controls the opacity of the slice from completely opaque to fully transparent. Emphasis on structures behind and in front of the slice can be changed by a slider, ranging from -1 to 1. At -1 the DVR behind the slice is emphasized and the opacity of the DVR in front is set to zero. At 0, both the sides of the slice have equal emphasis and at 1 only the DVR in front are displayed.

An example rendering of a synthetic dataset can be seen in Figure 5. On the left, regular volume rendering and slice rendering can be seen. Inserting the deformable plane in between the two spheres, it wraps around on either side. The illustrative slice rendering shows the slice blended with the local DVR. Volumetric clipping, on the other hand, completely removes one sphere while the other remains visible.

With regular clipping, understanding the shape of the cut is fairly easy, as the cut is flat. The shape of a deformed surface might be more difficult to interpret. Shading the surface of the cut can aid in understanding the shape of the membrane and aids in understanding how the structures in view spatially correlate.

Typically, shading in ray casting uses the gradients as surface normals. This does not work at the surface of the cut itself as the gradient might be parallel to the tangent of the

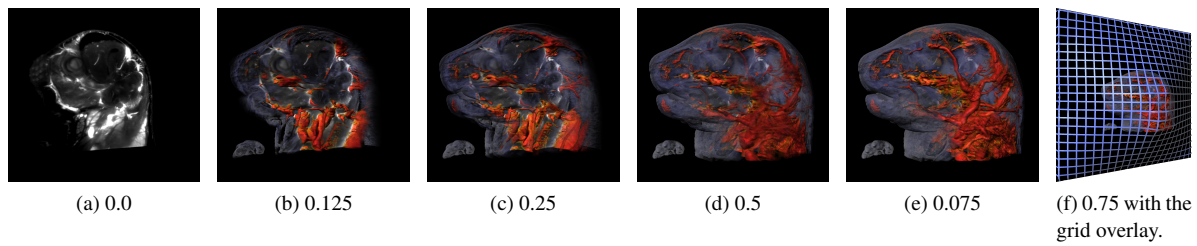


Figure 6: A mouse imaged using ultramicroscopy [DLS*07]. Adjusting the transparency according to the distance to the original slice. The figures show the effect of adjusting the relative maximum distance for contributing to the image.

cut surface. To shade the surface of the cut, we use normals of the membrane in the first hit of the volume rendering, similar to Weiskopf et al. [WEE03]. We render the normals of the deformed membrane into a separate texture which is then used during the ray casting.

4. Implementation

The prototype was implemented on a system with an Intel Core i7 3 GHz CPU and a GeForce GTX 580 graphics card, running Windows 7. Our goal was to create a technique which can handle reasonable sized data sets and run both the physics simulation and rendering at interactive frame rates. The physical simulation was more computationally demanding than the CPU could process. Therefore, we chose to implement our technique using the high performance GPU-based physics library *NVIDIA PhysX Engine* [phy11]. The library provided us with a cloth data structure which we could use as the elastic membrane for our technique, as well as data types for defining force kernels and collision elements in a physics environment constituting our force field.

5. Results

Illustrators have the option of selecting which structures they wish to keep in the illustration. In visualization of 3D data without any semantic information, this selection becomes very difficult. Clipping allows the user to spatially select which elements should be included. The membrane can be

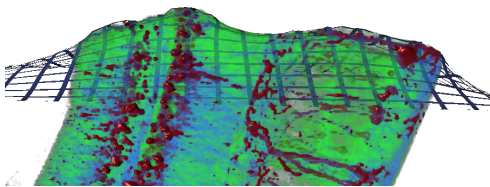


Figure 8: The hippocampus in a mouse brain imaged using ultramicroscopy [DLS*07]. Moving the slice plane through, the membrane is held back by the nerve cells (red), preserving it in the final image.

used to selectively clip seismic datasets. Here, the horizons below the sea bed are of high interest to the geo-scientific domain. However, seismic data has a high presence of noise and is difficult to explore using standard DVR and clipping. Applying a clipping surface adapting to the data attributes reveals more of the horizon compared to a flat clipping plane, as shown in Figure 7. Here we used only the PFTF to create force kernels from high intensity samples (red). The membrane is then moved smoothly over the volume. Other types of datasets have structures completely surrounded by similar data values. In these situations, creating the membrane inside the data might provide a more suitable option. In Figure 8, the slice wraps around the nerve cells in the hippocampus of a mouse brain. The elongated structures are then preserved in the final image.

Basic slice viewing is the most common technique for exploring large 3D data. With a combination of a slice rendering and membrane defined local DVR, we can get a better view of the structure close to the slice, such as the blood vessels and bone in Figure 3. The additional information can provide a better contextual perception. These images were created by drawing a simple sketch on a slice of the main aorta, which then automatically generated the PFTF. 3D ultrasound data is difficult to interpret using DVR due to much noise and similar data values. Local DVR rendering combined with slice rendering, can be a powerful tool for selective display of data and can aid in spatial understanding of dense 3D ultrasound volumes. In Figure 9, the elongated blood vessel structures are preserved as 3D structures in the image. Capturing the blood vessels required some time to define the correct PFTF, due to many data samples with very similar values.

6. Discussion

Our illustrative clipping system is composed of two main elements, i.e., the potential field and the elastic membrane. Their behaviour and their interactions are determined by the physics-based simulator, and the whole process depends on a fair number of parameters which have to be set manually. The optimal configuration is admittedly highly dependent on

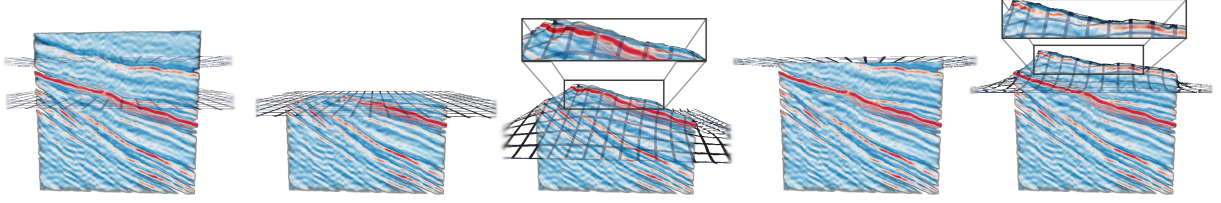


Figure 7: Revealing the structure of a seismic horizon in seismic amplitude data. Flat clipping compared to illustrative clipping.

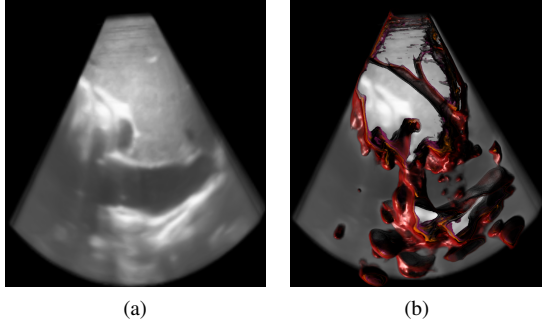


Figure 9: Certain structures are completely lost using basic slice viewing of 3D ultrasound (9a). Using the illustrative slicing technique, elongated structures such as the blood vessels are kept in the image (9b).

the purpose of the visualization scenario. We have explored the usage of the various parameters for both the potential field and the membrane, and here we give some guidelines to help the user to setup the system according to his/her needs.

One of the most relevant parameters is the scaling factor c , in Equation 3, which affects the strength of the force fields. Its value is determined through the PFTF, which has to be set according to scalar values in the dataset under consideration. In general, weak force fields make the membrane behave like a traditional clipping plane, while strong forces better emphasize inner structures, but may cause large oscillations of the membrane's position. All the examples in the paper have been generated with values of c up to 1000 and the membrane has never shown significant instability.

The global structure of the potential field is mainly affected by the positions and the areas of influence of every kernel. The placement is controlled by the octree level parameter: low level values produce dense kernel distributions, which, in turn, make the detection of the structures of interest highly accurate. The hierarchical ordering of voxels allows handling of even large datasets at interactive frame rates; however, in this case, generating a membrane which *wraps* around small structures can be difficult, as details may be lost due to averaging. In practice, the octree level can be used to control the tradeoff between speed and accuracy, and

it has to be set, taking into account the dataset's size and resolution, and the available computational power. Figure 10 shows the effect of using two different levels.

The area of influence of the kernels depends on the parameter a_j . Since the membrane should not intersect structures of interest, this value has to be set so that adjacent force fields overlap. On the other hand, the larger a_j is, the less tightly the membrane will fit to the structures of interest. After extensive testing, we found out that the best results are achieved when adjacent kernels overlap for 50% of their volume, i.e. when $a_j = 2v_j$, with v_j given by the super-voxel size in the j^{th} direction.

Taking now into consideration the elastic membrane, the main concerns are the bending and stretching stiffness parameters, which determine the membrane's resistance to bending and stretching, respectively. In general, low values of these parameters make the clipping plane fit more tightly to the structures in the volume, while values close to 1 make the clipping plane smoother.

The final shape of the clipping plane is heavily affected by the structure of the mesh used to represent the membrane. It is important to set the mesh resolution high enough in order to capture the smallest structures of interest, but, as for the potential field resolution, too high values could compromise the interactivity of the approach. Moreover, increasing the resolution indirectly makes the membrane stiffer, therefore the kernel's strength c should be increased as well. We set an initial resolution of 50×50 , but we let users freely modify this value in order to match the minimum feature size of specific datasets.

Lastly, the PBD algorithm assumes that every particle of the membrane has a certain mass. This value does not need to resemble the actual mass of the simulated piece of cloth, it is simply used to control the inertia of the particles. In the implementation of PBD provided by the NVIDIA PhysX framework, this parameter is optional, so we have not specified it. Its influence on the resulting visualization will be the subject of future investigation. The parameter settings for the examples in the paper are summarized in Table 1. Firstly, in all the cases, the default mesh resolution lead to fairly detailed results with no serious impact over the frame rate, so, unless there are very specific requirements, the average user won't need to modify it. The torso example can be used to

Data (res.)	Hierarchy	# of kernels	FPS w/physics	FPS display only	Figure
Torso ($256 \times 256 \times 556$)	3	2,3k	9	15	Figure 3 and 10b
Mouse ($424 \times 279 \times 190$)	2	15k	3	12	Figure 6
Seismic ($121 \times 121 \times 121$)	2	16k	5	13	Figure 7
Hippocampus ($212 \times 191 \times 44$)	1	16k	3	5	Figure 8
3D Ultrasound ($181 \times 245 \times 190$)	3	5k	10	14	Figure 9
Torso ($256 \times 256 \times 556$)	4	300	12	15	Figure 10a

Table 1: Performance results from the technique. The numbers are taken from the images which are included in this paper. FPS w/physics shows the FPS while the physics simulation is running, and FPS display only shows when it is paused. Resolution of the membrane for all images was set to 50×50 .

compare the effects of different octree levels: a higher value results in higher frame rates, with some minor differences in the rendered images (see the comparison in Figure 10).

The final results were shown to domain experts in the geological, medical and molecular biology domain. Geoscience experts described a recent work case where several horizons containing high-amplitude anomalies were dipping. The only way to display these surfaces were either to take a time slice (horizontal slice that does not consider the structural dip), or interpret a horizon and then create a surface that is used as a basis for making horizon slices. The second approach would be accurate, but also extremely time-consuming. Our method might be very useful to get a quick overview of high-amplitude anomalies and their extents, with an attractive trade-off between processing time and accuracy.

Medical experts stated that our deformable clipping approach could be interesting to combine with haptics for detecting cirrhosis, where dense *knots* form in the tissue and are otherwise difficult to detect in ordinary scanning. Overall, their comments stated that deformable clipping is promising for quickly gaining an overview of structures hidden in dense data.

Additionally, results from the adaptive seismic dataset clipping were shown to a molecular biologist and artist who is the author of the illustration in Figure 1. Currently, he is using automated methods for stylized rendering of entire molecular structures, but when it comes to revealing internal structures he needs to turn to a hand-craft. After we explained that the same approach can be utilized for adaptive clipping in molecular biology as well, he stated that our method could possibly automate the process of selective rendering, which is only possible nowadays by hand-drawing, and could aid in creating highly expressive images.

7. Summary and Conclusion

In this paper, we have presented an illustrative clipping technique for volume exploration using a deformable clipping geometry. By using a deformable geometry, we have demonstrated how we can utilize physically simulated soft bodies

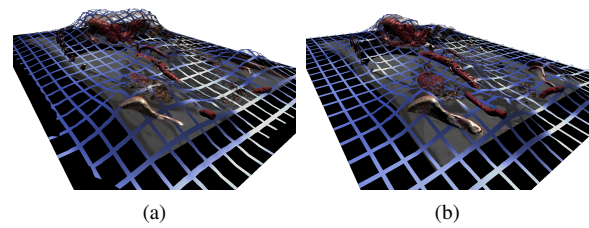


Figure 10: Using a higher traversal depth in 10a for the data hierarchy results in a less close fit compared to 10b.

to create expressive visualizations of volumetric data. The data attributes are translated into a potential field through a simple transfer function, defining how the data interacts with an elastic membrane. In addition, we have shown how the elastic membrane can be utilized in volume rendering, by aiding in removing visual clutter and adding context in well established volume exploration tools. Using a hierarchical structure and the GPU based physics library, interactive frame rates are achieved for both solving the physics and visualizing the data.

Acknowledgements

This work has been carried out within the IllustraSound research project (# 193170), which is funded by the VERDIKT program of the Norwegian Research Council with support of the MedViz network in Bergen, Norway (PK1760-5897-Project 11). In addition, this work has been partially funded by the Austrian Science Fund (FWF) through the ViMaL project (grant no. P21695) and by the Petromaks program of the Norwegian Research Council through the Geoillustrator project (#200512). We would like to thank our medical partners, Odd Helge Gilja and Trygve Hausken, as well as geologist Sten-Andreas Grundvåg, for valuable input. Finally, we would like to thank Laura Nic Lochlainn for aiding in proof reading and improving the final manuscript, Statoil for providing the seismic dataset, and David S. Goodsell for the motivating molecular machines illustration.

References

- [BG06] BRUCKNER S., GRÖLLER M. E.: Exploded Views for Volume Data. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (9 2006), 1077–1084. 2
- [BHW*07] BURNS M., HAIDACHER M., WEIN W., VIOLA I., GRÖLLER E.: Feature Emphasis and Contextual Cutaways for Multimodal Medical Visualization. In *Proceedings of EUROGRAPHICS* (2007). 2
- [BHWB07] BEYER J., HADWIGER M., WOLFSBERGER S., BUHLER K.: High-Quality Multimodal Volume Rendering for Preoperative Planning of Neurosurgical Interventions. *IEEE Transactions on Visualization and Computer Graphics* (2007). 2
- [DLS*07] DODT H.-U., LEISCHNER U., SCHIERLOH A., MAUCH N. J. C. P., DEININGER K., DEUSSING J. M., EDER M., ZIEGLGÄNSBERGER W., AND K. B.: Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods* 4, 4 (March 2007), 331–336. 7
- [FWH09] FUCHS R., WELKER V., HORNEGGER J.: Non-convex polyhedral volume of interest selection. *Journal of Computerized Medical Imaging and Graphics* 34, 2 (2009), 105–113. 2
- [goo12] Illustration by David S. Goodsell, the Scripps Research Institute. <http://mgl.scripps.edu/people/goodsell>, 2012. 2
- [HB00] HOUSE D. H., BREEN D. E. (Eds.): *Cloth modeling and animation*. 2000. 3
- [HNW93] HAIRER E., NØRSETT S. P., WANNER G.: *Solving Ordinary Differential Equations I*, 2nd ed. Springer Series in Computational Mathematics. Springer, Berlin Heidelberg, 1993. 4
- [ISC07] ISLAM S., SILVER D., CHEN M.: Volume Splitting and Its Applications. *IEEE Transactions on Visualization and Computer Graphics* 13, 2 (2007), 193–203. 2
- [KFW*02] KANITSAR A., FLEISCHMANN D., WEGENKITTL R., FELKEL P., GRÖLLER M. E.: *CPR - Curved Planar Ref-ormation*. Tech. rep., Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11/186, A-1040 Vienna, Austria, 2002. human contact: technical-report@cg.tuwien.ac.at. 2
- [Kre99] KRETZTECHNIK AG: 3D Ultrasound: A Dedicated System. *European Radiology* 9 (1999), S331–S333. 2
- [KVLP04] KONRAD-VERSE O., LITTMANN A., PREIM B.: Virtual Resection with a Deformable Cutting Plane. In *SimVis* (2004), pp. 203–214. 2
- [MHHR07] MÜLLER M., HEIDELBERGER B., HENNIX M., RATCLIFF J.: Position Based Dynamics. *Journal of Visual Communication and Image Representation* 18, 2 (2007), 109 – 118. 3, 4
- [MTCK*04] MAGNENAT-THALMANN N., CORDIER F., KECKEISEN M., KIMMERLE S., KLEIN R., MESETH J.: Simulation of Clothes for Real-time Applications. In *Proceedings of EUROGRAPHICS* (2004). 3
- [NMK*06] NEALEN A., MÜLLER M., KEISER R., BOXERMAN E., CARLSON M.: Physically Based Deformable Models in Computer Graphics. *Computer Graphics Forum* 25, 4 (2006), 809–836. 3
- [NWKY05] NAKAO M., WATANABE T., KURODA T., YOSHIMURA H.: Interactive 3D Region Extraction of Volume Data Using Deformable Boundary Object. In *Medicine Meets Virtual Reality 13: The Magical Next Becomes the Medical Now*. IOS Press, 2005, pp. 349–359. 2
- [phy11] Nvidia PhysX Engine. <http://uk.geforce.com/hardware/technology/physx>, 2011. 7
- [SSA*08] SCHULTZ T., SAUBER N., ANWANDER A., THEISEL H., SEIDEL H.-P.: Virtual Klingler Dissection: Putting Fibers into Context. *Computer Graphics Forum* 27, 3 (May 2008), 1063–1070. 2
- [TMS*06] TIETJEN C., MEYER B., SCHLECHTWEG S., PREIM B., HERTEL I., STRAUSS G.: Enhancing Slice-based Visualizations of Medical Volume Data. In *Proceedings of the Eurographics / IEEE VGTC Symposium on Visualization* (2006), pp. 123–130. 2
- [VKG04] VIOLA I., KANITSAR A., GROLLER M. E.: Importance-Driven Volume Rendering. In *Proceedings of IEEE Visualization* (2004), pp. 139–146. 2
- [WEE02] WEISKOPF D., ENGEL K., ERTL T.: Volume clipping via per-fragment operations in texture-based volume visualization. In *Proceedings of IEEE Visualization* (2002), pp. 93–100. 2
- [WEE03] WEISKOPF D., ENGEL K., ERTL T.: Interactive Clipping techniques for Texture-Based Volume Visualization and Volume Shading. *IEEE Transactions on Visualization and Computer Graphics* 9, 3 (2003), 298 – 312. 7

2.3 Continuous Levels-of-Detail and Visual Abstraction for Seamless Molecular Visualization

Július Parúlek, Daniel Jönsson, Timo Ropinski, Stefan Bruckner, Anders Ynnerman, and Ivan Viola

Computer Graphics Forum, Volume 33, Issue 6, pages 276–287, 2014

DOI: 10.1111/cgf.12349

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Continuous Levels-of-Detail and Visual Abstraction for Seamless Molecular Visualization

Julius Parulek¹, Daniel Jönsson², Timo Ropinski², Stefan Bruckner¹, Anders Ynnerman² and Ivan Viola^{1,3}

¹Department of Informatics, University of Bergen, Bergen, Norway
{julius.parulek, stefan.bruckner}@uib.no

²Department of Science and Technology, Linköping University, Linköping, Sweden
{daniel.jonsson, timo.ropinski, anders.ynnerman}@liu.se

³The Institute of Computer Graphics and Algorithms, Vienna University of Technology, Vienna, Austria
viola@cg.tuwien.ac.at

Abstract

Molecular visualization is often challenged with rendering of large molecular structures in real time. We introduce a novel approach that enables us to show even large protein complexes. Our method is based on the level-of-detail concept, where we exploit three different abstractions combined in one visualization. Firstly, molecular surface abstraction exploits three different surfaces, solvent-excluded surface (SES), Gaussian kernels and van der Waals spheres, combined as one surface by linear interpolation. Secondly, we introduce three shading abstraction levels and a method for creating seamless transitions between these representations. The SES representation with full shading and added contours stands in focus while on the other side a sphere representation of a cluster of atoms with constant shading and without contours provide the context. Thirdly, we propose a hierarchical abstraction based on a set of clusters formed on molecular atoms. All three abstraction models are driven by one importance function classifying the scene into the near-, mid- and far-field. Moreover, we introduce a methodology to render the entire molecule directly using the A-buffer technique, which further improves the performance. The rendering performance is evaluated on series of molecules of varying atom counts.

Keywords: level of detail algorithms, implicit surfaces, clustering, scientific visualization

ACM CCS: Computer Applications [J.3]: Life and Medical Sciences Biology and Genetics COMPUTER GRAPHICS [I.3.3]: Picture/Image Generation Viewing algorithms

1. Introduction

Molecular visualization today is challenged by molecular dynamics (MD) simulations with the requirement of displaying huge amounts of atoms at interactive frame rates for the visual analysis of binding sites. Simulated data sets do no longer consist of only one moderately sized macromolecule, but instead of molecular systems representing complex interactions, for example, a phospholipid vesicle membrane together with proteins anchored in the membrane (Figure 1, right). One can easily obtain data sets where tens or hundreds of thousands of atoms are animated throughout a series of 1000 time-steps.

To analyse a binding site, a special visual representation is most popular among molecular biologists known as the solvent-excluded surface (SES) [Ric77]. This representation directly conveys information whether a solvent of a certain size is able to reach a particular binding site on the surface of the macromolecule. While this repre-

sentation is valued by the molecular biology domain, it is also expensive to compute. To achieve interactivity with the scene, biologists sacrifice information provided by SES and investigate molecules with blobby Gauss kernel representations [Bli82], or with a simple space filling approach. The latter one, for example, can be represented very quickly by impostor-based sphere splatting, but it does not answer precisely whether a solvent can bind at a specific location to a macromolecule. Another research question is how to abstract the molecular surface further, beyond representing each single atom. An example of where such an abstraction would be essential is a *Powers-of-Ten* zooming interactive environment where the user can zoom-in onto a single atom, or zoom out to see a cellular-level structure of the same entity. At the cellular-level it is completely out of the question to render each atom of a molecule and a hierarchical abstraction would be needed. In the search for the appropriate solution we turn to the visual crafts for inspiration, which have been already successfully applied on molecular visualization [vdZLB11].

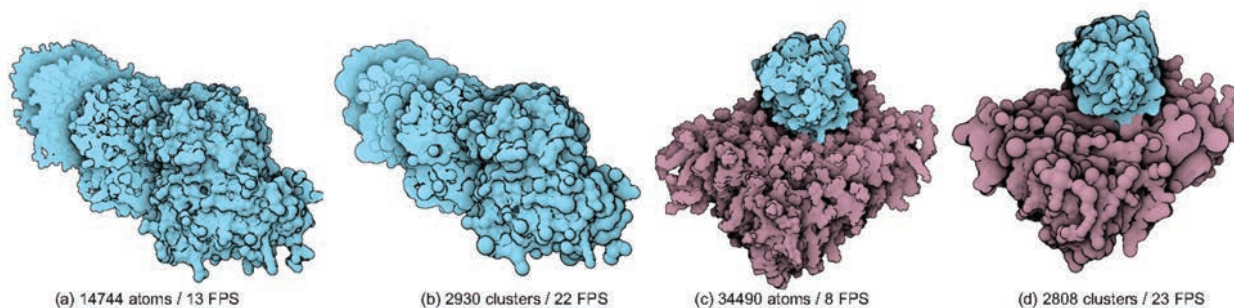


Figure 1: Two molecular examples, Tubulin and phospholipase bound to lipid membrane, demonstrating utilization of our seamless visual abstraction. We employ three different surface representations [solvent-excluded surface (SES), Gaussian kernels and van der Waals spheres], their corresponding shading abstractions (diffuse shading and contours, constant shading with contours, constant shading without contours) and hierarchical representation. The application of individual levels is based on the distance to the camera; that is, the closest surface is based on highest surface, shading and hierarchical levels while the farthest are displayed via the lowest ones. In the presented examples we achieved $5\times$ to $10\times$ speed-up as compared to the full SES representation (a and c), and $10\times$ to $20\times$ when additionally applying hierarchical abstraction (b and d).



Figure 2: Object constancy employed in visual arts by Winsor McCay 'When Black Death Rode'.

Illustrators sometimes take a different approach to visually abstracting molecules from details. Instead of modifying the molecular representation into an entirely different molecular abstraction (such as transition between space filling representation and ribbons showing β -sheets and α -helices), they effectively use the perceptual principles of object constancy to depict structures that are too far away to recognize the details, through a simplified representation of that object. In this way the illustrators' manual creation process is speedup while at the same time also resulting in a more convenient visualization for the viewer, whose cognitive processing related to object constancy autocompletes the simplified visual representation with an object instance. A beautiful utilization of this approach can be seen on Winsor McCay's artwork of 'When Black Death Rode' shown in Figure 2, which was exemplified by the professional scientific illustrator Bill Andrews [And06].

To address the molecular visualization challenge delineated above we propose to employ a seamless level-of-detail rendering scheme, in the same way as illustrators approach rendering of scenes containing multiple instances of the same object, and taking advantage of the object constancy perceptual principle. As a general rule, closest

to the viewer we aim at providing a maximum of relevant information related to the structure and binding sites. We also utilize the level of detail scheme to guide the viewer to relevant information in the spirit of focus + context visualization techniques. The most detailed molecular surface representation is the SES representation, where every atom (except hydrogens) is rendered to form the molecular surface. Farther away from the viewer, we smoothly change the visual representation to an approximation of SES through Gaussian kernels. Structures farthest away from the view are represented by simple sphere splatting. When the individual atoms are no longer discernible, we employ hierarchical clusterings, where the atoms at a particular spatial location are grouped into *super-atoms*, which enables lower memory requirements and faster rendering performance. The use of these three levels of detail are motivated by the cognitive zones of the viewer: focus, focus-relevant and context zone (Figure 1). Generalizing the concept leads us to the definition of a 3-D importance function that can be based on the distance measure from a molecular feature, not only as a distance from the camera.

Nevertheless, the question that remains unanswered is how we can preserve smoothness in detail-level transitions. Smoothness in transitions is an important requirement as an abrupt change in level-of-detail will become a salient artefact that will involuntarily attract the attention of the biologist. To tackle this problem, we propose to utilize an implicit surface representation, where we can seamlessly blend from one surface representation to another one. The seamless illustration-inspired level-of-detail scheme for molecular systems based on implicit surfaces is the main contribution of this paper. Additionally, the scheme fulfills the focus and context model, where both levels are blended via the seamless transformations. While illustrative representations have been investigated in the context of molecular visualization earlier, they have never been investigated within the context of a level-of-detail scheme.

The contributions of this paper can be summarized as follows: We propose a novel visualization approach that increases the overall rendering performance by utilizing a level-of-detail concept applied via hierarchical abstraction, surface abstraction and shading

abstraction. We build upon our earlier work [PRV13] on seamless molecular abstraction and extend it with respect to several aspects. Most notably, we present a method for hierarchical abstraction, which goes beyond the level of atomic detail. For this purpose we hierarchically cluster the entire molecular structure at various detail levels.

2. Related Work

Our approach builds on several aspects of previous work on molecular visualization, in particular with respect to choosing appropriate visual representations, methods for interactive rendering and level-of-detail techniques.

2.1. Visual representations

Tarini *et al.* present a real-time algorithm for visualizing molecules with the goal to improve depth perception [TCM06]. By combining ambient occlusion and edge-cueing together with graphics processing unit (GPU) data structures, they achieve interactive frame rates for molecules of up to the order of 10^6 atoms. Based on this representation, the authors report an improved understanding of the molecule structure. While we exploit different representations mainly in order to allow for efficient rendering, Lueks *et al.* combine different representations of a molecule in a single view in order to support understanding of different abstraction levels [LVvdZ*11]. By allowing the user to control the seamless transition between different molecule representations, these can be viewed in a combined manner and thus reveal information at different degrees of structural abstraction. The abstractions which are combined are based on previous work presented by van der Zwan *et al.* [vdZLBI11]. The authors classify molecular representations based on their *illustrativeness*, structural abstraction and spatial perception. By giving the user control over these three parameters, s/he can change the depiction of a molecule. Thus the possible representations largely resemble known molecular representations widely used in text books. The *illustrativeness* presented by van der Zwan *et al.* is achieved by combining different rendering styles. Similar to the work done by Tarini *et al.* [TCM06], they also experiment with ambient occlusion techniques. In contrast, Weber presents a cartoon-style rendering algorithm for protein molecules, which exploits GPU shaders to generate interactive pen-and-ink effects [Web09]. In the work of Cipriano and Gleicher [CG07], spatio-physico-chemical properties are used to generate a simplified representation that conveys the overall shape. This approach, like many of the presented illustration models, goes back to the original work done by David Goodsell [Goo09], who has developed a simplistic, but expressive style for representing molecules through space filling. His approach combines ambient occlusion with cel-shading and silhouettes in order to illustrate residuals. This illustration approach has for instance been recently adopted by Falk *et al.* [FKE12], and it also inspired the creation of the renderings shown in this paper.

2.2. Interactive rendering

Besides recent efforts dealing with the visual representation of molecules, a lot of work has been dedicated to increase the over-

all rendering performance. For instance, Sharma *et al.* present an octree-based approach, which allows billions of atoms to be rendered interactively by exploiting view-frustum culling [SKNV04]. A combination of probabilistic and depth-based occlusion algorithms is used during rendering to determine the visible atoms. More recently, Grottel *et al.* have investigated different data simplification strategies which also incorporate culling [GRDE10]. In particular, they take into account data quantization, video memory-based caching and a two-level occlusion culling strategy. Lampe *et al.* focus on the visualization of slow dynamics for large protein assemblies [DVRH07]. To represent these large-scale dynamic models, they also use a hierarchical approach where the topmost layer represents residues as the high-level building blocks of a molecule. For each residue only orientation information is sent to the GPU, where the generation of the individual atoms is performed on-the-fly. Since SES represents the most advanced representation of molecular surfaces, which allows the molecule interactions and evolution to be studied, some effort has also been dedicated to improving the rendering of these fairly complex structures. Parulek and Viola propose an SES representation based on implicit surfaces [PV12]. By exploiting constructive solid geometry operations on these surfaces, they obtain implicit functions which locally describe a molecule's surface. As their ray-casting-based rendering of this representation requires no pre-processing, they are able to vary SES parameters interactively. Frey *et al.* focus on MD simulation data [FSG*11]. In order to speed up rendering of these data, they reduce the amount of particles by focusing on those considered as relevant for the visualization. In contrast to our technique this resembles a data reduction approach instead of a data simplification approach.

2.3. Level-of-detail techniques

Level-of-detail approaches have a long history in computer graphics [LWC*02]. Most techniques for rendering molecular data use surface simplification methods for generating different LODs [KOK06]. Lee *et al.* [LPK06] visualize large-scale molecular models using an adaptive level of detail (LOD) technique based on a bounding tree. Fraedrich *et al.* [FAW10] sample only the visible particles in the scene into perspective non-uniform grids in view space. These optimizations result in low computation times even for large data sets. They render the isosurfaces using GPU-based ray-casting. Krone *et al.* [KSES12] use a view-independent volumetric density map representation and generate a surface representation for rendering using a GPU implementation of the Marching Cubes algorithm similar to the work of Dias *et al.* [DG11]. The work of Bajaj *et al.* [BDST04] incorporates a biochemically sensitive level-of-detail hierarchy into the molecular representation and uses an image-based rendering approach. Although presented in the context of visual data mining in document collections, the H-BLOB method by Sprenger *et al.* [SBG00] is of relevance as it uses a hierarchical clustering and visualization approach based on implicit surfaces. Our approach maintains an implicit representation throughout the pipeline and uses it for rendering directly. We use a hierarchical data representation scheme which forms, together with visual representation, and surface representation, a 3-D abstraction space. This provides us with fine-grained control over the different representational dimensions and enables us to flexibly

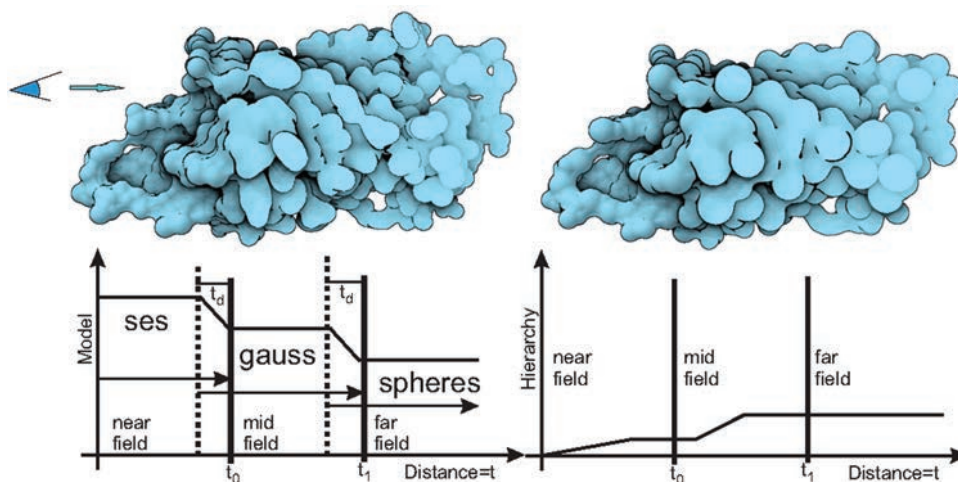


Figure 3: Left: The organization of the three surface and shading levels according to importance function $t(\mathbf{p})$ defined by the increasing distance from the camera. In the overlapping zones, the representations are merged using linear interpolation. The molecule is displayed with the full atom count, 1852 atoms. Right: The extraction of cluster hierarchies based on $t(\mathbf{p})$. As the distance from the camera increases, the clusterings are retrieved from the higher hierarchical levels representing bigger clusters. An illustration shows exploitation of hierarchical abstraction containing 784 clusters, that is, 42% compress ratio.

and seamlessly adjust the level of abstraction during interactive visualization.

3. Methodology

Motivated by the need for visualization of large molecular systems, we propose a seamless visual abstraction scheme which provides continuous transitions from the computationally expensive, but most relevant visualization technique, to the fastest representation which is suitable for representing the context. The key component of our approach which enables this seamless transition is an implicit surface representation on which all the visual abstractions are based on. We define three different levels of visual abstraction, with overlapping transition zones: a near-field, a mid-field and a far-field. The field boundaries are defined by an importance function, $t(\mathbf{p})$. Besides the distance from the viewer used as our primary example, the importance function can be thought of as a distance measure from an interesting molecular feature (e.g. a cavity) or from a region of interest, interactively specified by the user (e.g mouse cursor location) [PRV13]. Our LOD visual abstraction consists of three distinct categories: hierarchical abstraction, surface abstraction and shading abstraction.

The first level of abstraction is concerned with whether the molecule is represented directly by atoms, or whether the atoms are grouped into clusters, that is, *superatoms*, which are then represented by a ball of a larger radius covering the volume of the grouped atoms (Figure 3, right). Our previous work [PRV13] included only the atomic level. Section 4.3 discusses the generation of this hierarchy.

The second category is the visual abstraction of surfaces. The most domain-relevant visual representation is the SES. Based on this representation the molecular biologists can decide whether

a specific binding site is accessible to a solvent or not. The intermediate visual abstraction level is based on a Gaussian kernel representation that approximates the SES and is often used in analysis of molecular surfaces despite its lower expressive value with respect to the binding sites [KFR11]. This visual abstraction is a compromise between rendering performance and expressiveness. The last level of the proposed visual abstraction scheme is a space-filling approach where individual atoms are represented by spheres. This is the fastest representation to render, however, its main usefulness is in providing a more gross structural context rather than providing a useful information about a local molecular detail (Figure 3).

The third category is concerned with the visual abstraction of shading. Together with geometry, we abstract the details in shading in the following way. For conveying shape detail, we employ a local diffuse shading model. For conveying relative depth, ambient occlusion is used. Ordinal depth cues are communicated with contour rendering and the figure-ground ambiguity is resolved by silhouette rendering. This scheme is motivated by the workflow that David Goodsell, an acknowledged molecular scientist and illustrator, employs in molecular illustrations [Goo09]. We additionally provide a detail level with local shading. While Goodsell's illustrations have equal amount of visual cues for the entire molecular system, we have a specific distribution of visual cues for each level of detail. The figure-ground separation, which uses silhouette and ambient occlusion as a relative depth cue, is used for all abstraction levels. The near- and mid-field levels additionally convey structural occlusion with contour rendering as an ordinal depth cue. The near-field conveys the shape and therefore uses diffuse shading, while the other two levels are represented with a constant shading, abstracting from atomic details. An example incorporating all abstraction levels is shown in Figure 3. The overall molecular rendering is performed by means of a ray-casting method where each ray is incrementally

processed, thereby allowing us to evaluate corresponding molecular and shading models.

4. Molecular Visual Abstraction

The main reason for choosing an implicit representation is that it enables us to easily form a smooth transition, or a blend, between different types of surfaces. For instance, when two implicit functions f and g overlap in space, a simple way to generate a seamless transition between them is via linear interpolation: $h = (1 - t)f + tg$. This preserves the continuity of even two different representations, which is a necessary property in order to achieve a seamless transition between different molecular models. This property would be very hard to achieve with any boundary representation, especially on a real-time basis. We propose a set of abstraction levels which are aligned with visual processing, but our framework can also easily handle additional levels. In our work the interpolation parameter t is interpreted as an importance value $t = t(\mathbf{p})$ that varies with the position \mathbf{p} in the scene. In our demonstrations, we use the distance from the camera as the importance function, $t(\mathbf{p}) = \|\mathbf{eye} - \mathbf{p}\|$. We specify borders for all three areas (near-, mid- and far-field) using $t(\mathbf{p}) \leq t_0 \equiv$ near-field, $t_0 < t(\mathbf{p}) \leq t_1 \equiv$ mid-field and $t(\mathbf{p}) > t_1 \equiv$ far-field. The length of the transition area is controlled by t_d , which defines the blending interval between two distinct molecular surface representations. Thus, when a point \mathbf{p} lies only in one area we can evaluate a single implicit function, while for the overlapping areas we need to evaluate both functions and combine their result by linear interpolation.

4.1. Surface abstraction

We assume a set of atoms defined as $C = \{(\mathbf{c}_1, r_1), \dots, (\mathbf{c}_n, r_n)\}$ and introduce the three implicit functions each defining the molecular model for one of the three intervals.

4.1.1. SES representation

To represent SES by the means of implicits, we take as a basis the approach proposed by Parulek and Viola [PV12]. The method for evaluating the implicit function has cubic complexity $O(n^3)$. The final implicit function evaluates an exact Euclidean distance to the surface, although only to the distance R from the iso-surface of SES representation. One of the advantages of the proposed method is the flexibility of varying the parameters during rendering, for example, atoms participating in SES representation or the solvent radius R . This is the main reason why this method is incorporated into our pipeline; it enables us to vary the length of the near-field easily. This representation is the one that is most computationally expensive and it makes sense to apply it only when studying inter-atomic cavities in detail. Therefore, although in principle applicable to all hierarchical abstraction levels, it is only meaningful to utilize it in the near-field focal region.

4.1.2. Gaussian kernel representation

For the second, mid-field level of surface abstraction, we utilize the Gaussian model, which is widely used as an approximation of

the SES model. It smoothly blends the density field generated by the atoms and also forms a seamless transition between the SES and sphere models. The utilization of the Gaussian kernel for implicit modelling was used for the first time by Blinn [Bli82] to describe the electron density function of atoms by summing the contribution from each atom as follows: $F_{gauss}(\mathbf{p}) = T - \sum_i b_i e^{-a_i d_i^2}$, where d_i represents the distance from \mathbf{p} to the centre of atom c_i , b_i represents the *blobbiness*, a_i describes the atom radius and T defines the electron density threshold. We adopted Blinn's model and specified the parameters a_i and b_i as were introduced by Grant and Pickup [GP95]: $b_i = R^2$, $a_i = -\ln r_i^2 / 2b_i$ and $T = 0.5$.

4.1.3. van der Waals sphere representation

Let us define a set of implicit functions defined as $\{f_1, f_2, \dots, f_n\}$, where each $f_i(\mathbf{p}) = r_i - \|\mathbf{p} - \mathbf{c}_i\|$ represents an atom \mathbf{c}_i with the corresponding van der Waals radius r_i . The implicit function defining the union of spheres can be written as $F_{spheres}(\mathbf{p}) = \max\{f_1(\mathbf{p}), f_2(\mathbf{p}), \dots, f_n(\mathbf{p})\}$, where the maximum operator represents the union term [Ric72]. In order to render the iso-surface of $F_{spheres}$ solely, we actually do not need to evaluate the intersection of the ray and the function by a root finding method. Instead, rendering can efficiently be performed by ray-casting the spheres directly and storing the closest depth values to the camera in a depth buffer. Therefore, even if the function evaluation still has $O(n)$ complexity, the entire rendering pipeline can be optimized by drawing all the spheres in parallel, while the atomic operations evaluate the depth buffer. Moreover, the rendering performance can be increased by utilizing the sphere billboard technique [DVRH07]. To form a smooth blend between the van der Waals spheres and the Gaussian kernels representation we only need to evaluate $F_{spheres}$ in the transition area $t(\mathbf{p}) \in [t_1 - t_d, t_1]$. The sphere billboarding technique is employed when $t(\mathbf{p}) > t_1$.

We utilize linear interpolation between the representations inside transition zones, while the remaining zones only require a single representation to be evaluated. Our approach allows all three level-of-detail areas and their lengths to be modified in real time. We choose linear interpolation as it represents a simple, intuitive and efficient solution. More sophisticated approaches, for example, variational methods [TO99] or extended space mapping [SP98] provide several parameters to fine-tune the shape of the final interpolation, but are relatively expensive to evaluate and not yet suitable for real-time rendering applications.

4.2. Shading abstraction

Our shading model employs a set of visual abstractions that selectively enhance shape and depth information. The shading scheme is inspired by the approach presented by David Goodsell's artwork [Goo09]. We use his system of visual cues, that is, constant shading, contour and depth enhancement, which he employs in molecular illustrations, although applied only on our sphere representation. We apply these visual cues in a focus and context manner, where the focus is represented for the interval $t(\mathbf{p}) < t_0$. In the remainder of this section, we discuss the application of the aforementioned visual cues according to all three level-of-detail areas.

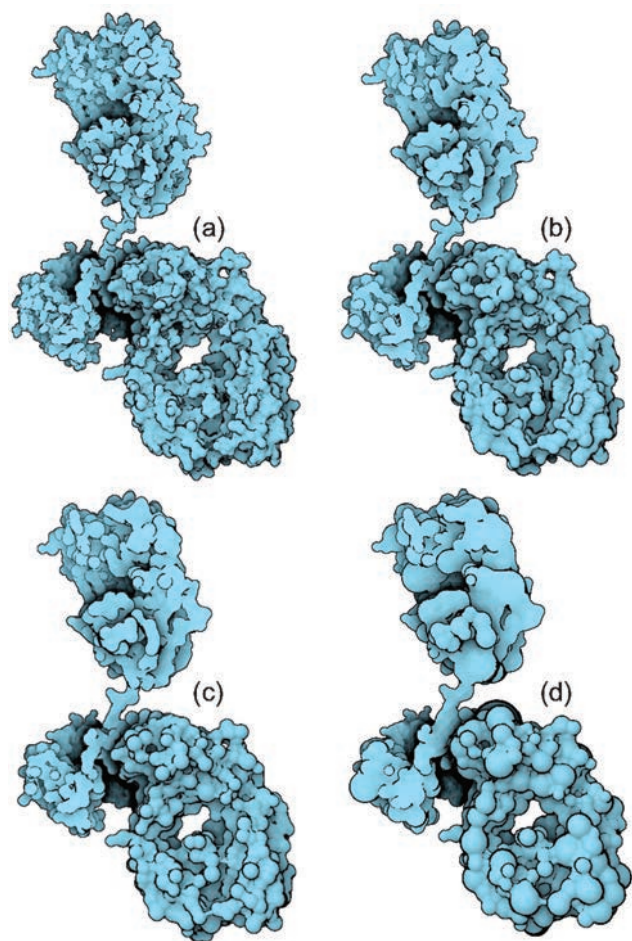


Figure 4: Four hierarchical levels on immunoglobulin. The first level represented by the full atom count (a), 12 530 atoms. The second level (b), approximates the atoms in the first level by a set of 6990 clusters, which represents 55.7% elements of the full atom count. The other levels, (c) and (d), contain 3122 (24.9%) and 1576 (12.5%) elements.

In the near-field $t(\mathbf{p}) \in [0, t_0]$, we employ a local diffuse shading model (DM) in combination with the constant shading model (CM), which is applied in accordance with the $t(\mathbf{p})$ value. This enables us to create much smoother transitions to CM. In the transition zone $t(\mathbf{p}) \in (t_0 - t_d, t_0]$, we interpolate the shading model such that the DM continuously disappears towards the end of the transition area.

In the mid-field and far-field zones, $t(\mathbf{p}) \in [t_0, \infty)$, we employ constant shading model. The reason for applying the CM in the mid-field is that the Gaussian model conveys lower accuracy for the solvent shape than the SES. Thus, by using CM we are able to visually decrease the surface discrepancies between the two models (Figure 1). Besides the shading we incorporate silhouettes and contours into our visualization. We employ the approach of Kindlmann *et al.* [KWTM03] to generate contours of uniform thickness using the fast view-dependent curvature approximation of Krüger *et al.* [KSW06]. Furthermore, we preserve the contours

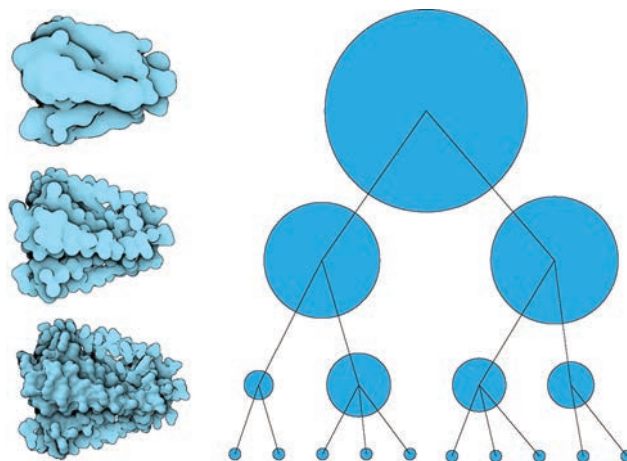


Figure 5: A hierarchy based on spatial clustering is created using a bottom-up approach. The hierarchy is combined with the surface and shading abstractions using a seamless transition model.

for near- and mid-field but neglect it in the far-field. The reason behind discarding the contours in the context area defined by spheres is that they do not fully emphasize the inter-spherical space, that is, just enhancing the spherical shape. In the second transition area, we scale the *contour* predicate to make the contour disappear continuously.

The silhouettes are generated with respect to the background of the rendered molecule, that is, all the pixels that do not belong to the molecule are considered background. Afterwards, in image space, we perform edge detection on the binary texture where 1 represents molecule and 0 means background. The silhouette is preserved for all three zones. This was chosen to imitate the Goodsell's approach and, additionally, to enhance the overall shape of the molecule. As the last step in our rendering pipeline we add screen space ambient occlusion based on the method proposed by Luft *et al.* [LCD06]. The ambient occlusion is, similarly to the silhouettes, applied to all three zones.

4.3. Hierarchical abstraction

The visual representations discussed so far are based on access to the original resolution of the whole data set represented by individual atoms. Unfortunately, this limits the visualization to data sets that can fit and be streamed into memory in a timely manner. These representations are necessary when a structure is explored in detail, that is, by being able to go down to the atom level. While this is often desirable for the structures being directly in focus, or near the focus field, structures being far outside the focus do not need to convey this detailed information. Still, it is important that the overall structure of the molecule is conveyed and that large-scale features are preserved. Moreover, since state-of-the-art molecular visualization techniques [KBE09, LBPH10, PV12] exploit almost instant surface generation from the initial set of atoms prior to rendering the surface, it is necessary to enhance the rendering of large molecular and even cellular scenes using a new data simplification scheme as well.

To achieve this new level of abstraction, we again drew inspiration from David Goodsell's representations. By condensing the visualized information to the most essential visual elements, he is able to communicate the important features without introducing additional clutter. As the dominant visual elements in his representations are silhouettes and surfaces having a flat appearance, we have designed our hierarchical abstraction such that we can reduce a molecule to these elements. One alternative to achieve such an abstraction would be to use primary or secondary structures analysis. However, this would not generalize to other kinds of molecules, for example, lipids. Therefore, we propose to exploit location-based clustering, which enables support for a wider spectrum of molecules while generating a hierarchy of nested surface representations. By using a location-based clustering these nested surface representations will have a high degree of conservation with respect to the outline of the molecule (Figure 4). In the following, the initial set of atoms, $C = \{(\mathbf{c}_1, r_1), \dots, (\mathbf{c}_n, r_n)\}$, becomes a set of spherical elements. Such an element can be either an atom or a cluster described again by a centre \mathbf{c}_i and a radius r_i . More importantly, the function evaluation procedure is the same for all three surface representations. While the presented approach is generalizable, it can also be integrated with the implicit surface abstractions in order to create a seamless visual transition for the user.

In order to form the hierarchical representation of the particles, we use a bottom-up approach. The hierarchy generation process therefore starts with performing a spatial clustering on the original atom data. After the clustering is complete each formed cluster is represented by a sphere with a radius r , which bounds all the particles within the cluster, and the cluster centre c , which is the centre of gravity computed from cluster members. The next level of detail is created using the clusters from the previous level of detail as input and raising the error threshold by a factor of two. This process continues until a maximum number of levels of details have been created or only a single cluster remains. Each particle/cluster will therefore have a single parent, but can have multiple children as illustrated in Figure 5.

Several clustering methods have been studied, where the computational complexity was considered crucial. We analysed the following clustering algorithms: DB-SCAN, k-means, hierarchical and the Affinity Propagation (AP) technique [Llo82, FD07, Mü13]. As a notable result, we found out that applying a density-based clustering scheme does not perform well on molecular data sets. The main reason is that molecular objects inherently lack any significant variation within the atom density distribution. Therefore, applying the DB-SCAN often produces only a single cluster for the entire molecule. Moreover, by testing various molecules, it became clear that qualitatively the AP algorithm performed best for the tested data sets. The AP algorithm was presented by Frey and Dueck [FD07], having formed clusters more uniformly and with a lower error bound than any other clustering algorithms. However, the biggest drawbacks of the AP algorithm is its computation complexity, $O(n^2)$. Indeed, AP is by far the slowest algorithm in our test group. Even though the cluster coverage between neighbouring levels is far better than using the remaining techniques, when the atom count is more than 10 K, creating already a single cluster level takes tens of minutes. Similarly, k-means also provides a computationally expensive so-

lution, which prohibits its application on large molecules. On the other hand, the fast hierarchical clustering method, proposed by Müllner [Mü13], offers a very fast solution with a high-quality cluster coverage at the same time. A comparison of applying the three clustering algorithms is depicted in Figure 6. Generation of five hierarchical levels takes 30 923 (ms) for AP, 23 508 (ms) for k-means and 676 (ms) for hierarchical clustering.

5. Rendering and Performance Analysis

Our rendering pipeline consists of several steps (Figure 7). In the first one, we traverse the cluster hierarchy in top-down manner to retrieve all the clusters/atoms that are about to be used for the molecular representation and visualization. Starting from the highest level, we evaluate whether a cluster C is directly used in the visualization or whether it is required to recursively evaluate its child nodes (clusters or atoms in the leaves). The evaluation criterion that decides whether a cluster C is going to be added to the display list is defined by function

$$g(C, t) \equiv t > \frac{l_C}{l_{max}} t_1, \quad (1)$$

where l_C is the hierarchy level of C , l_{max} is the highest available hierarchy level and t_1 represents the far-field depth. When a cluster meets the criteria, it is added to the display list (Figure 8). The hierarchy traversal is performed on the CPU side, as a frame pre-processing, before the display list is sent to GPU. Here we have not found any performance drop even for larger hierarchical trees.

In the second step, we render clusters/atoms, stored in the display list, as spheres with an increased cluster radius that defines their area of influence. This area is defined by means of solvent diameter $2R$, that is, each cluster is rendered as a sphere with its cluster radius increased by $2R$. The reasoning why to choose the solvent diameter as an area of the atom influence is described by Varshney *et al.* [VBW94]. Moreover, we do not perform sphere ray-casting, but instead quickly splat spheres using billboarding [TCM06].

Instead of displaying these spheres, we store them in the so-called A-buffer. The theoretical framework describing the A-buffer was presented by Carpenter in 1984 [Car84]. Essentially, A-buffer is a linked list of fragments generated for every pixel separately using atomic operations on the GPU. We define one global atomic counter that serves as the head pointer to the linked list. This counter is increased by one every time when a new fragment is generated in the fragment shader. Each fragment record consists of the entry and the exit depth of a rendered cluster, and the cluster id. The fragment record is then stored in the shared image at the location addressed by the global counter. It is noteworthy to mention that similar approaches for rendering molecules, defined by blobby objects and iterative blending, were presented by Szecs and Illes [SI12] and Parulek and Brambilla [PB13], respectively.

In the third step, before the actual ray-casting, we sort the fragment records by entry depth. This allows us to easily step along those clusters during subsequent ray-casting. Sorting is performed using CUDA, as it proved to be substantially faster (more than a factor of 4) than a fragment shader implementation in our experiments. Thus

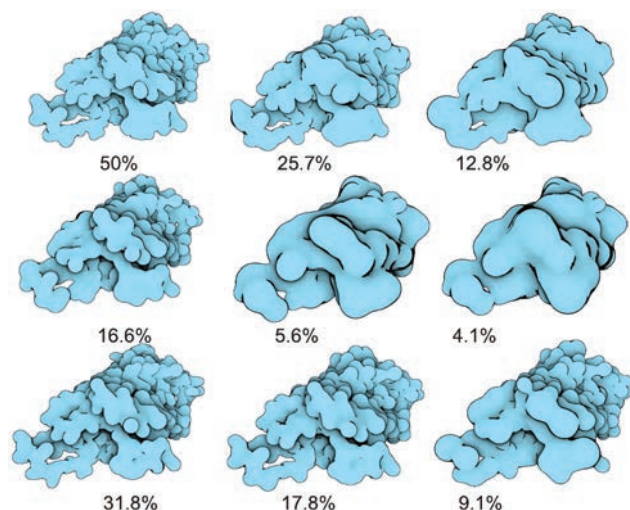


Figure 6: Comparison of three clustering algorithms. The top row presents the first three levels of k -means clustering, the second row demonstrates affinity propagation method and the third row stands for fast hierarchical clustering. The percentage represents the ratio between the cluster render list size and the full atom count. Note that the ratios are different due to characteristics of the employed algorithms providing dissimilar clusters.

for each image pixel (ray), we obtain a list of clusters that influence the function evaluation along the ray in ascending order.

In the fourth step, the scene is rendered. Here the ray is cast for each image pixel, where we generate an input 3-D point \mathbf{p} based on the entry depth of the first sphere at the pixel location and the projection matrix. Afterwards, we employ a sphere tracing algorithm [Har94] that processes the ray in a stepwise fashion until the last sphere exit depth is reached or we hit the iso-surface, that is, $|F| \leq \epsilon$. The selection of ϵ can be used to either increase the surface detail or to improve the rendering performance. When a point on the ray is in the area where no sphere of influence is presented, the point is automatically shifted to the first unprocessed sphere along the ray, that is, the next one in the linked list. This allows us to perform empty space skipping very efficiently.

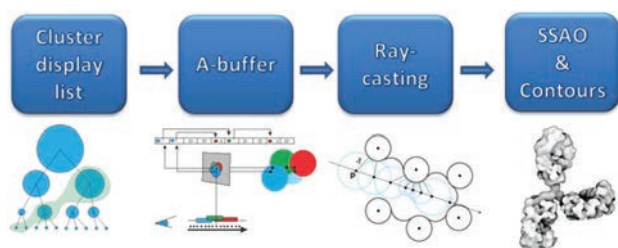


Figure 7: An illustration of the rendering pipeline. The formation of the cluster display list is determined by Equation (1). Clusters are represented as spheres that are rendered into the A-buffer. The ray-casting is performed through sphere tracing algorithm. In the end, we compute screen space ambient occlusion.

Here we describe the performance analysis, where the lengths of individual fields across the molecule are varied. We show that the user has the possibility to alter the fields to either get more molecular details with decreased frames per second (FPSs) or vice versa.

While a comprehensive evaluation of the performance of our method with respect to all parameter combinations (varying lengths of all three fields, the length of the transition area and iso-surface precision parameter ϵ) is not feasible, we demonstrate its performance using several indicative examples. For the hierarchical abstraction, we utilize the fact that the performance is linearly dependent on the amount of clusters used. For instance, in Figure 1, when using just 20% (b) and 8% (d) of spheres compared to full atom count, (a) and (c), the performance increases almost $2\times$ and $3\times$. Therefore, we focus our performance analysis rather on the surface abstraction when using the full atom count. We introduce evaluation based on several examples of molecules of various sizes where we alter the lengths of near-, mid- and far-field, while choosing a fixed size for the transition area as well as the precision parameter. We setup $t_d = 4R$ and $\epsilon = 0.05R$, where R is the solvent radius. The performance measurements are performed on a workstation equipped with two (2 GHz) processors and 12.0 GB RAM and with the GPU, NVIDIA GeForce GTX 690.

It is important to mention that for each frame we perform all the steps presented in Section 5. One of the biggest advantages of our real-time implicit function evaluation is the possibility of varying the function parameters anywhere in space, while preserving an interactive system response. To generate a suitable description of the performance based on the lengths of three fields, we store all FPS values for each distribution of fields. Afterwards, we employ ternary plots displaying a coverage of the three areas in barycentric coordinates. The colours, from yellow to red, encode the achieved FPS. For simplicity, we use relative length of fields expressed in percentage of how much of the molecule participates to each field;

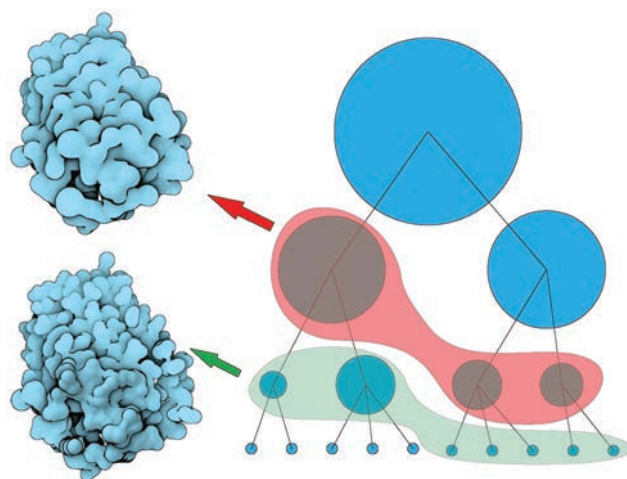


Figure 8: An example of two display lists. During the cluster tree traversal, all the nodes that fulfill Equation (1), are added to the list. By zooming out outwards the molecule, the red display list becomes more reduced and abstracted than the green display list.

for example, $t_0 = 1/3$ and $t_1 = 2/3$ represents equally distributed fields over the molecule, which is represented by the central point in all four plots. This evaluation method is applied to four molecules (Figure 9), Aquaporin (1852 atoms) (a), proliferatic cell nuclear antigen (12 555 atoms) (b), phospholipase bound the lipid membrane (34 490 atoms) (c), asymmetric chaperonin complex (58 674 atoms) (d).

6. Results and Limitations

We demonstrate our technique on several molecules of various sizes. We employ the Protein Data Bank (PDB) file format, which stores the molecular information and atom positions.

A typical demonstration of our technique is when the lengths of fields vary over the molecule and we fix the fields boundaries t_0 and t_1 and perform interactive zoom-in towards the molecular centre. Such an example is displayed in Figure 10 using the full atom count and also the hierarchical representation. Notice that on each zoom level there are some visual differences, but the higher cluster levels apply only when a molecule moves away from the viewer, where the visual discrepancies are even more suppressed. On the other side, in this example and for comparative purposes the non-clustered and clustered versions are depicted in the same size. In the figure, only 60% of spheres were employed for the rightmost visualization and 30% of spheres for

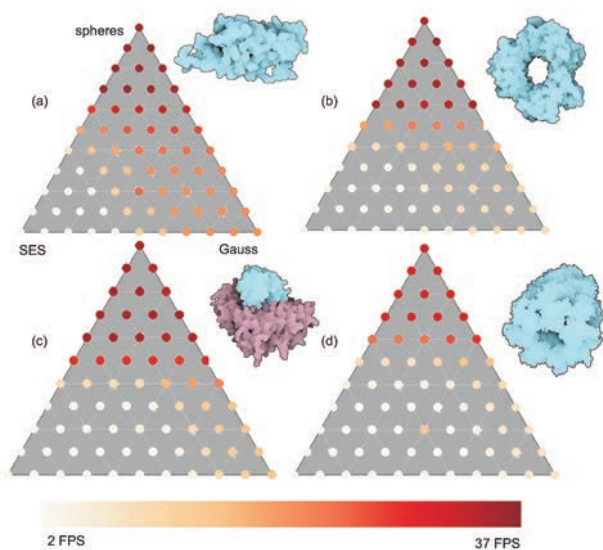


Figure 9: Ternary plots showing performance analysis evaluated on four distinct molecular structures. The analysis is based on the lengths of individual fields (SES, near-field; Gauss, mid-field; spheres, far-field). (a) Water channel (Aquaporin). (b) Proliferatic cell nuclear antigen. (c) Phospholipase bound the lipid membrane. (d) Asymmetric chaperonin complex. Note that the achieved FPS are, in the case of the camera-based importance function, directly proportional to the lengths of each areas; that is, prolongation of the near-field leads to decreasing FPSs on the other side, contraction of the far-field increases FPSs.

the leftmost compared to the spheres/atoms contained within the molecule.

Through our LOD concept we are able to boost the rendering performance of molecular models by $5\times$ to $10\times$, while keeping the most detailed SES representation for the closest parts of the molecule from the camera. Additionally, when applying the hierarchical representation, we get even up to $20\times$ the frame rate compared to full SES representation. All three surface representations are evaluated on-the-fly during ray-casting, which provides us with a great flexibility with regards to either enhancing the performance or the details for dynamic data sets.

The utilization of hierarchical abstraction brings two major limitations. The first one is the actual surface precision when using the full atom count compared to exploiting the cluster hierarchy. Here, our shading abstraction helps to hide the most of the surface dissimilarities (Figure 10). Nevertheless, to compute the error quantitatively, we would need to firstly evaluate the most suitable parameters for the clustering method, for example, distance metric and stopping criteria, to reduce the error there first. The cluster error increases as we move up in the hierarchy. Nevertheless, the highest levels are usually employed only when depicting contextual molecular parts being farther away from the viewer.

The second limitation of utilizing the hierarchical abstraction is the requirement of performing the sequential clusterings. This has to be done for each new structure modification repetitively. For molecules containing a few thousands of atoms, formation of 5 to 10 hierarchical clusterings can take up to 1 s. While for larger molecules (molecular systems) this can take up to a minute. For example, forming five levels for the lipid–protein complex (Figure 1) took 20 s, while generation of five levels for asymmetric chaperonin complex (58 674 atoms) takes 80 s (Figure 11).

Nevertheless, we can already see the potential of our approach for visualizing mesoscopic whole-cell simulations [FKE12]. Here a cluster hierarchy can be formed in the pre-computation step for all acting molecules. Another potential solution to perform a clustering on dynamic structures, is to exploit a fast GPU-based bounding volume hierarchies (BVH). For instance, Bitner *et al.* introduce a GPU-based solution to update a BVH tree to minimize the overall cost function [BHH13], which in our case can be based on one of the abstraction levels.

We have demonstrated our method to biologists and a scientific illustrator, where we acquired a feedback about the overall visual quality and possible extensions of the proposed technique. Firstly, the illustrator was pleased with the results and the originality of the proposed concept. On the other hand, it was suggested to improve the contour rendering for the SES portion of the model. Here the main issue he raised was that the contours can appear jaggy which is due to C^1 discontinuities on the iso-surface of the SES model. Such discontinuous areas are also hard to track via the sphere tracing algorithm, which we also employ for the contour predicate. Additionally, the problem may be amplified by the fast curvature approximation we employ and a more costly scheme could help to overcome it. Overall, however, these issues were not seen as critical.

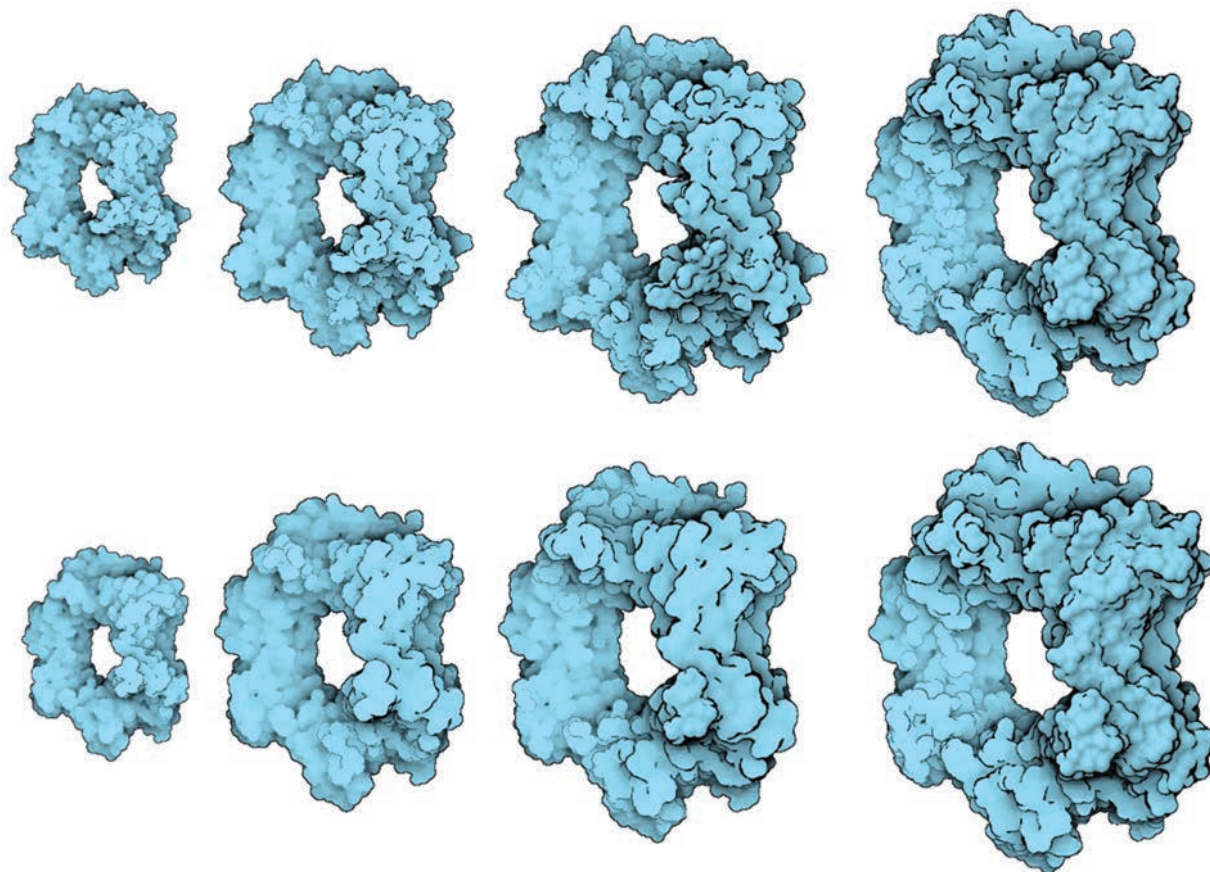


Figure 10: Comparison of zooming in towards the molecule (proliferative cell nuclear antigen) performed using the full atom count (12555 atoms, top) and the hierarchical representation (bottom). The display list contains from left to right: 3716, 4527, 5292 and 7516 clusters.

Furthermore, we were suggested to incorporate additional silhouettes into the final visualization to clearly delineate boundaries between distinct molecules in compound systems. While not the focus of this paper, we found that this is an important note to be considered in our future work. Domain experts found the achieved visuals original and helpful, mainly due to the interplay between the visualizations and the precision. Furthermore, they suggested to apply the proposed method to more application-oriented scenarios.

7. Summary

We have proposed a novel approach for visualization of molecular surfaces. Our approach is capable of rendering large protein complexes interactively, while rapidly reducing the amount of displayed primitives, and at the same, keeping the visual appearance similar to the original data. Our method utilizes the level-of-detail concept by means of three different molecular surface models, SES, Gaussian kernels and van der Waals spheres combined in one visualization. Moreover, we introduced three shading levels that are aligned with

the three surface models. For the realization, we took an inspiration from illustrations showing densely populated scenes with similar objects (spheres model with almost no detail), which are smoothly interconnected with highly detailed structures (SES model with full details) through the visual abstraction (Gaussian kernels model with fading out details). Finally, we proposed a new hierarchical abstraction that approximates the molecular atoms with a set of clusters that are employed in the final visualization.

The importance function that represents the choice of the surface, shading and hierarchical models is based on the distance from the camera. We showcased how this can be used effectively to increase the rendering performance, even for large molecules, by interactive specification of level-of-detail boundaries. The entire rendering pipeline is performed on-the-fly. We introduced an LOD shading scheme with respect to all three fields individually. We preserved a seamless transition of depth, figure and shape visual cues using interpolation of shading and model schemes. A figure-ground ambiguity is solved via the utilization of the silhouette. The silhouette also keeps the entire molecule, even divided into distinct fields, perceptually unified.

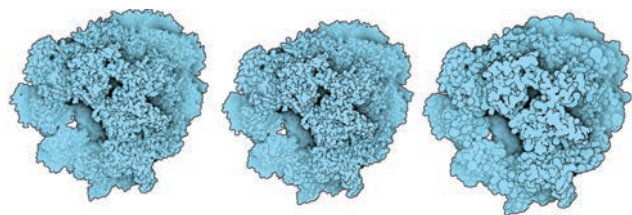


Figure 11: An example of asymmetric chaperonin complex (58 674 atoms). Our representation can easily depict large molecules, where we employ the full atom count (left), 36 740 (middle) and 9989 (right) clusters.

Acknowledgements

We thank Nathalie Reuter for providing the MD simulation data sets, David Goodsell and Helwig Hauser for giving us the necessary feedback for the overall visualization. This work has been carried out within the PhysioIllustration research project (# 218023), which is funded by the Norwegian Research Council. This paper has been also supported by the Vienna Science and Technology Fund (WWTF) through project VRG11-010, and supported by the EC Marie Curie Career Integration Grant through project PCIG13-GA-2013-618680, and also by grants from the Excellence Center at Linköping and Lund in Information Technology (ELLIIT), the Swedish Research Council through the Linnaeus Center for Control, Autonomy, and Decisionmaking in Complex Systems (CADICS), and the Swedish e-Science Research Centre (SeRC), as well as VR grant 2011-4113.

References

- [And06] ANDREWS B.: Introduction to ‘perceptual principles in medical illustration’. In *ACM SIGGRAPH 2006 Courses* (New York, NY, USA, 2006), SIGGRAPH ’06, ACM.
- [BDST04] BAJAJ C., DJEU P., SIDDAVANAHALLI V., THANE A.: Texmol: Interactive visual exploration of large flexible multi-component molecular complexes. In *Proceedings of IEEE Visualization 2004* (Austin, TX, USA, 2004), pp. 243–250.
- [BHH13] BITTNER J., HAPALA M., HAVRAN V.: Fast insertion-based optimization of bounding volume hierarchies. *Computer Graphics Forum* 32, 1 (2013), 85–100.
- [Bli82] BLINN J.: A generalization of algebraic surface drawing. *ACM Transactions on Graphics* 1 (1982), 235–256.
- [Car84] CARPENTER L.: The a -buffer, an antialiased hidden surface method. *SIGGRAPH Computer Graphics* 18, 3 (January 1984), 103–108.
- [CG07] CIPRIANO G., GLEICHER M.: Molecular surface abstraction. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1608–1615.
- [DG11] DIAS S. E., GOMES A. J.: Graphics processing unit-based triangulations of blinn molecular surfaces. *Concurrency and Computation: Practice and Experience* 23, 17 (2011), 2280–2291.
- [DVRH07] DAAE LAMPE O., VIOLA I., REUTER N., HAUSER H.: Two-level approach to efficient visualization of protein dynamics. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1616–1623.
- [FAW10] FRAEDRICH R., AUER S., WESTERMANN R.: Efficient high-quality volume rendering of SPH data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1533–1540.
- [FD07] FREY B. J., DUECK D.: Clustering by passing messages between data points. *Science* 315 (2007), 972–976.
- [FKE12] FALK M., KRONE M., ERTL T.: Atomistic visualization of mesoscopic whole-cell simulations. In *EG Workshop on Visual Computing for Biology and Medicine* (Norrköping, Sweden, 2012).
- [FSG*11] FREY S., SCHLOMER T., GROTTTEL S., DACHSBACHER C., DEUSSEN O., ERTL T.: Loose capacity-constrained representatives for the qualitative visual analysis in molecular dynamics. In *IEEE Pacific Visualization Symposium* (Hong Kong, China, 2011), pp. 51–58.
- [Goo09] GOODSSELL D.: *The Machinery of Life*. Springer, New York, 2009.
- [GP95] GRANT J. A., PICKUP B. T.: A Gaussian description of molecular shape. *The Journal of Physical Chemistry* 99, 11 (March 1995), 3503–3510.
- [GRDE10] GROTTTEL S., REINA G., DACHSBACHER C., ERTL T.: Coherent culling and shading for large molecular dynamics visualization. *Computer Graphics Forum* 29, 3 (2010), 953–962.
- [Har94] HART J. C.: Sphere tracing: A geometric method for the antialiased ray tracing of implicit surfaces. *The Visual Computer* 12 (1994), 527–545.
- [KBE09] KRONE M., BIDMON K., ERTL T.: Interactive visualization of molecular surface dynamics. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1391–1398.
- [KFR11] KRONE M., FALK M., REHM S.: Interactive exploration of protein cavities. *Computer Graphics Forum* 30, 3 (2011), 673–682.
- [KOK06] KANAI T., OHTAKE Y., KASE K.: Hierarchical error-driven approximation of implicit surfaces from polygonal meshes. In *Proceedings of the Eurographics Symposium on Geometry Processing 2006* (Sardinia, Italy, 2006), pp. 21–30.
- [KSES12] KRONE M., STONE J., ERTL T., SCHULTEN K.: Fast visualization of Gaussian density surfaces for molecular dynamics and particle system trajectories. In *Proceedings of EuroVis 2012 Short Papers* (Vienna, Austria, 2012), pp. 67–71.
- [KSW06] KRUGER J., SCHNEIDER J., WESTERMANN R.: Clearview: An interactive context preserving hotspot visualization technique. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (September–October 2006), 941–948.

- [KWTM03] KINDLMANN G., WHITAKER R., TASDIZEN T., MOLLER T.: Curvature-based transfer functions for direct volume rendering: Methods and applications. In *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)* (Washington, DC, USA, 2003), VIS '03, IEEE Computer Society.
- [LBPH10] LINDOW N., BAUM D., PROHASKA S., HEGE H.-C.: Accelerated visualization of dynamic molecular surfaces. *Computer Graphics Forum* 29, 3 (2010), 943–952.
- [LCD06] LUFT T., COLDITZ C., DEUSSEN O.: Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics* 25, 3 (July 2006), 1206–1213.
- [Llo82] LLOYD S.: Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28, 2 (1982), 129–137.
- [LPK06] LEE J., PARK S., KIM J.-I.: View-dependent rendering of large-scale molecular models using level of detail. In *Proceedings of the International Conference on Hybrid Information Technology 2006* (Cheju Island, Korea, 2006), pp. 691–698.
- [LVvdZ*11] LUEKS W., VIOLA I., VAN DER ZWAN M., BEKKER H., ISENBERG T.: Spatially continuous change of abstraction in molecular visualization. In *Abstracts of 1st IEEE Symposium on Biological Data Visualization* (BioVis 2011, Providence, RI, USA, 2011), M. Meyer and C. Nielsen (Eds.), IEEE Computer Society.
- [LWC*02] LUEBKE D., WATSON B., COHEN J. D., REDDY M., VARSHNEY A.: *Level of Detail for 3D Graphics*. Elsevier Science Inc., New York, NY, 2002.
- [Mül13] MÜLLNER D.: fastcluster: Fast hierarchical, agglomerative clustering routines for r and python. *Journal of Statistical Software* 53, 9 (5 2013), 1–18.
- [PB13] PARULEK J., BRAMBILLA A.: Fast blending scheme for molecular surface representation. *IEEE Transactions on Visualization and Computer Graphics* 19 (December 2013), 2653–2662.
- [PRV13] PARULEK J., ROPINSKI T., VIOLA I.: Seamless abstraction of molecular surfaces. In *Proceedings of the 29th Spring Conference on Computer Graphics* (Comenius University, Bratislava, Slovakia, 2013), pp. 120–127.
- [PV12] PARULEK J., VIOLA I.: Implicit representation of molecular surfaces. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis 2012)* (Songdo, Korea, March 2012), pp. 217–224.
- [Ric72] RICCI A.: A constructive geometry for computer graphics. *The Computer Journal* 16, 2 (1972), 157–160.
- [Ric77] RICHARDS F. M.: Areas, volumes, packing, and protein structure. *Annual Review of Biophysics and Bioengineering* 6, 1 (1977), 151–176.
- [SBG00] SPRENGER T. C., BRUNELLA R., GROSS M.: H-blob: A hierarchical visual clustering method using implicit surfaces. In *Proceedings of IEEE Visualization 2000* (Salt Lake City, Utah, USA, 2000), pp. 61–68.
- [SI12] SZECSEI L., ILLES D.: *Real-Time Metaball Ray Casting with Fragment Lists* (2012). C. Andujar and E. Puppo (Eds.), Eurographics Association, Cagliari, Sardinia, Italy, pp. 93–96.
- [SKNV04] SHARMA A., KALIA R. K., NAKANO A., VASHISHTA P.: Scalable and portable visualization of large atomistic datasets. *Computer Physics Communications* 163, 1 (2004), 53–64.
- [SP98] SAVCHENKO V., PASKO A.: Transformation of functionally defined shapes by extended space mappings. *The Visual Computer* 14, 5–6 (1998), 257–270.
- [TCM06] TARINI M., CIGNONI P., MONTANI C.: Ambient occlusion and edge cueing for enhancing real time molecular visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 1237–1244.
- [TO99] TURK G., O'BRIEN J. F.: Shape transformation using variational implicit functions. *Computer Graphics* 33, Annual Conference Series (1999), 335–342.
- [VBW94] VARSHNEY A., BROOKS JR, F. P., WRIGHT W. V.: Computing smooth molecular surfaces. *IEEE Computer Graphics Applications* 14 (September 1994), 19–25.
- [vdZLBI11] VAN DER ZWAN M., LUEKS W., BEKKER H., ISENBERG T.: Illustrative molecular visualization with continuous abstraction. *Computer Graphics Forum* 30, 3 (2011), 683–690.
- [Web09] WEBER J. R.: ProteinShader: Illustrative rendering of macromolecules. *BMC Structural Biology* 9, 1 (2009), 1–19.

2.4 Attractive Flicker: Guiding Attention in Dynamic Narrative Visualizations

Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, Ivan Viola

IEEE Transactions on Visualization and Computer Graphics, Volume 20, Issue 12, pages 2456-2465, 2014
DOI: 10.1109/TVCG.2014.2346352

©2014 IEEE. Reprinted, with permission, from Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, and Ivan Viola, Attractive Flicker: Guiding Attention in Dynamic Narrative Visualizations, IEEE Transactions on Visualization and Computer Graphics, December 2014. In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of TU Wien's products or services. Internal or personal use of this material is permitted.

Attractive Flicker

Guiding Attention in Dynamic Narrative Visualizations

Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, *Member, IEEE*, and Ivan Viola

Abstract—

Focus+context techniques provide visual guidance in visualizations by giving strong visual prominence to elements of interest while the context is suppressed. However, finding a visual feature to enhance for the focus to pop out from its context in a large dynamic scene, while leading to minimal visual deformation and subjective disturbance, is challenging. This paper proposes Attractive Flicker, a novel technique for visual guidance in dynamic narrative visualizations. We first show that flicker is a strong visual attractor in the entire visual field, without distorting, suppressing, or adding any scene elements. The novel aspect of our Attractive Flicker technique is that it consists of two signal stages: The first “orientation stage” is a short but intensive flicker stimulus to attract the attention to elements of interest. Subsequently, the intensive flicker is reduced to a minimally disturbing luminance oscillation (“engagement stage”) as visual support to keep track of the focus elements. To find a good trade-off between attraction effectiveness and subjective annoyance caused by flicker, we conducted two perceptual studies to find suitable signal parameters. We showcase Attractive Flicker with the parameters obtained from the perceptual statistics in a study of molecular interactions. With Attractive Flicker, users were able to easily follow the narrative of the visualization on a large display, while the flickering of focus elements was not disturbing when observing the context.

Index Terms—Visual attention, flicker, narrative visualization

1 INTRODUCTION

Dynamic visualizations of complex phenomena or real-time events can be interactively explored by the user, but at the same time they tell a story. Examples are 3D visualizations of physiological processes, crowd simulations, multiple coordinated views of dynamically changing stock exchange data, or observation and collaboration interfaces in emergency response centers. What these examples have in common is that a lot of dynamically changing data has to be displayed concurrently. This is often achieved by using large displays, such as projection screens or multi-monitor setups. As a result, the user is confronted with a large amount of information and visual clutter. A major challenge in such a setting therefore is to guide the user’s attention to the current main actors of the story while not introducing too much distraction from the remaining content or causing subjective annoyance.

Focus+context techniques make “*uneven use of graphics resources [...] to visually discriminate data-parts in focus from their context*” [14] and thereby guide the user’s attention to elements of importance. Common approaches exaggerate or suppress visual features of scene elements, like hue [8], luminance [22, 7], sharpness [25], or size [10], to achieve visual guidance. However, in large cluttered scenes with dynamic elements – like those found in the examples mentioned above – modifying these features may not be sufficient to generate a popout effect to reliably attract the user’s attention. Especially if visualizations exceed the user’s parafoveal field of view, such static visual attraction techniques may fail without any additional guidance, like trails [16, 39] or a virtual searchlight [22].

In addition, many of the above mentioned techniques visually deform data entities and thereby increase the probability that the scene is misunderstood. For instance, fisheye distortions can lead to misjudgments of relative distances [12], and blurring of context elements

makes it hard to identify contextual details [48]. Our goal therefore is to achieve visual guidance with minimal deformation of the data to maintain scene understanding.

A class of attraction techniques that has been demonstrated to work well under various conditions are those adding dynamic changes to focus elements. We are particularly interested in attractors that change the visual appearance of focus elements over time, *i.e.*, make them flicker. While being highly effective even in dynamic scenes [34], flickering attractors in graphical user interfaces are also known to be a source of annoyance for users [11, 16].

In this paper, we explore the design space of flickering attractors to make elements of interest stand out in large, complex visualizations while adding minimal annoyance for the observer. To achieve this goal, we introduce a new class of flickering attractors: *Attractive Flicker*. What distinguishes Attractive Flicker from other visual guidance techniques is that it uses a two-stage signal: A short, but highly salient, *orientation stage* is followed by a smooth decay into a minimally disturbing *engagement stage*. We empirically determined the flicker parameters to effectively attract the user’s attention in the orientation stage, and to generate low annoyance in the engagement stage, in two perceptual pilot experiments on a large display showing a dynamic scene. Finally, we demonstrate that Attractive Flicker effectively guides the user through a narrative visualization while adding low distraction in a study of large-scale molecular interactions.

2 BACKGROUND

Attractive Flicker utilizes knowledge of basic visual attention research to provide effective visual guidance. Therefore, we first summarize basics from the visual perception and cognition point of view, followed by related work on attention guidance.

2.1 Visual Attention

When observing an image, there are several factors that influence where our attention is directed. *Feature integration theory* (FIT) [41] suggests that multiple feature dimensions of an image – such as color, orientation, spatial frequency, brightness, and movement direction – are rapidly processed automatically and in parallel. This *preattentive* processing ability enables observers to catch certain properties of a scene in less than 250 milliseconds, *i.e.*, without moving the eyes to scan the scene. For instance, simple search tasks, like determining whether there is a single red dot within a number of blue dots, can be conducted preattentively (see Healey and Enns [15] for an overview).

- Manuela Waldner, Mathieu Le Muzic, Matthias Bernhard, Werner Purgathofer, and Ivan Viola are with Vienna University of Technology. E-mail: {waldner|mathieu|Matthias.Bernhard|wp|viola}@cg.tuwien.ac.at
- Werner Purgathofer is with VRVis Research Center. E-mail: Purgathofer@VRVis.at.

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014; date of publication xx xxx 2014; date of current version xx xxx 2014.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

Only after this preattentive processing, objects can be identified by integrating the separate features with the support of focused attention. Experiments with implementations of bottom-up saliency models (e.g., [20, 19]) showed that motion and flicker feature maps were stronger predictors of users' saccade directions than color, intensity, and orientation contrasts [18, 31], indicating that dynamic features are stronger visual attractors than static ones.

It is well known that bottom-up saliency is by far not the only factor influencing attention in visual scenes. *Guided search* theory by Wolfe *et al.* [47] assumes that visual attention is guided by both, bottom-up and top-down information. Both, stimulus-driven bottom-up and expectation-driven top-down processing, contribute to a common *activation map*, guiding the observer's attention in the scene.

2.2 Attention Guidance

Despite the fact that top-down attention plays a major role in visual attention, the influence of bottom-up attention should not be neglected. In psychology and education research, *selective processing* describes the phenomenon that learners are mainly able to reproduce insights encoded by perceptually salient image regions [26]. Healey and Enns [15] argue that directing the users' attention to important elements in a visualization may strengthen their engagement and, as a result, increase their insight.

To increase bottom-up saliency in regions of interest, focus+context techniques enhance certain features of focus elements or suppress features of context elements [14]. For instance, common ways to highlight elements of interest are to use distinct hues for the focus elements (commonly used for brushing and linking [8]), to decrease the luminance of context elements [22, 7], blur the context [25], or generally reduce the amount of details for context regions by simplifying the rendering style [6]. Traditional focus+context techniques apply spatial distortions to dedicate more screen space to focus elements, while suppressing the context (e.g., *Fisheye* techniques [10]). A different approach is to add artificial visual cues to the scene to make a target pop out. Examples are text labels for the selected elements (e.g., in *Gap-minder* [1]) or trails leading the user's gaze to highlighted elements [16, 39]. Most of these examples apply substantial modifications to the scene. In many cases, this is desirable, as the attention guidance mechanism itself also serves as an aesthetic scene property (for instance, when using stylized rendering [6]). In other cases, a strong visual deformation of the context may lead to an undesirable misinterpretation of the scene.

Another challenge of visual attention guidance is to find an image feature that can be modified to create a reliable popout effect in complex scenes. As formalized by Rosenholtz [35], a focus element needs to be sufficiently different from the context with respect to the modified feature to generate a popout effect. One approach to achieve this goal is to use an additional output modality to draw the user's attention to a certain location, such as spatial audio [30, 40]. As a purely visual solution, researchers have suggested to iteratively tune static image features in regions of interest to increase the bottom-up saliency of these regions for videos [42] or volume visualizations [23]. A different approach is to pick a feature to enhance that is underrepresented in a scene. For instance, on desktop interfaces, "moticons" [4] are icons effectively drawing the user's attention by simple motions. In information visualization, oscillating movements have been used for filtering and brushing [3], to highlight connected nodes in subgraphs [44], but also to visualize multivariate data [17, 45]. It could be shown that dynamic targets are much easier to detect than static ones, especially with increasing distance from the users' fixation point [33, 4]. But what if the scene itself contains inherent motion?

2.3 Flicker Attractors

A low-level visual feature that is typically not prominently used in both, static and dynamic visualizations is flicker. Flicker has a couple of interesting properties that make it attractive for attention guidance in visualization. Pinto *et al.* [34] could show in an experiment that blinking targets can be effortlessly discriminated from moving distractors. When flickering in coherent phases, even differences between flicker

frequencies can be easily detected [17]. Hoffmann *et al.* [16] found that blinking window frames were more effective to guide the user's attention on large displays than static attractors like a red frame around the window. However, their results also showed that blinking window frames were considered the most annoying visual cue. In contrast, in an experiment by Bartram *et al.* [4], users found that a 1 Hz blinking of icons is the least disturbing – but also least effective – motion cue. In a similar experimental setting, Gluck *et al.* [11] discovered that a proxy icon following the cursor was more effective, but also more annoying than a ~ 3 Hz blinking icon. These experiments indicate that there seems to be a trade-off between effectiveness of an attractor and its perceived annoyance. However, there is little knowledge how the two different parameters of a flicker signal – namely amplitude and frequency – interact and affect detection rate and perceived annoyance. The goal of our work is to answer this research question and to design a new class of flicker attractors based on empirical evidence.

An alternative approach to overcome the annoyance introduced by flickering attractors was presented by Bailey *et al.* [2]. Their *Subtle Gaze Direction* (SGD) technique uses 10 Hz luminance or hue flicker signals to attract the user's attention to targets in the peripheral vision. To avoid annoyance, they use an eye tracker to detect the user's current fixation point and saccade direction. Whenever the user initiates a saccade towards the flickering target, the flicker is stopped. Studies showed that SGD could reliably guide the users' gaze towards flickering targets in static images [2], as well as a sequence of image regions in narrative paintings [29], without users being aware of the flickering. We pursue a slightly different goal than subtle attention guidance techniques like SGD [2] or saliency modulation of static image features [23, 42]. Instead of subtly attracting the users' attention without their conscious knowledge, we want to provide a clearly visible, yet minimally disturbing signal the user can actively follow, to understand the narrative of a visualization. The grand challenge is to design a minimally distracting attractor that works with complex, dynamic visualizations, and does not prevent the user from simply walking up and start using the system without hardware-induced constraints or time-consuming calibration processes.

3 REQUIREMENTS FOR VISUAL GUIDANCE

When addressing the problem of attention guidance in large, dynamic visualizations, we are confronted with a number of requirements that rule out many traditional focus+context techniques:

R1: *Visual guidance should not be limited to the (para)foveal visual field. It should also work on large displays:* Humans are color-blind in the peripheral vision [32], which implies that attractors based on chromatic contrast are unsuitable for large displays. In general, dynamic cues are considered more effective visual attractors when using large displays than static ones [4, 43].

R2: *Visual guidance should also work in complex and cluttered scenes with a lot of different colors and elements changing their visual appearance or location over time:* From a neurobiological point of view, flicker is encoded in a separate feature map and therefore is less likely to interfere with other features like hue or motion direction [19, 34], which are more frequently used in visualizations than flicker. However, little is known how well a flicker signal can attract a user's attention in the peripheral field of view when the visualization itself contains a large number of heterogeneous and dynamic elements. We will address this research question in a comparative pilot experiment (Section 4) and in our first perceptual study (Section 6).

R3: *To enable visual guidance, modifications to the scene – both to the attractors and the distractors – should be minimal:* When integrating over time, the flicker feature (e.g., the object's geometry or appearance) remains constant. Thereby, it does not persistently modify the scene in an indeterminable way such as, for instance, a continuous change of hue or saturation, but can still be clearly discriminated from the context.

R4: *Scene modifications should lead to minimal annoyance or irritation of the user:* Flicker has been shown to be more irritating than many static attractors [16], but less distracting than other motion cues [4]. In our second perceptual study, we will therefore explore the

trade-off between flicker effectiveness and annoyance by controlling its amplitude and frequency (Section 7).

R5: Visual guidance should be possible without expensive or restricting hardware: There are attention guidance techniques that presumably work well in large, complex visualizations with negligible modifications to the visual scene by using spatial audio through tracked headphones (e.g., [40]) or an eye tracker to switch off a dynamic attractor on demand [2]. However, these techniques require non-commodity and potentially restricting hardware equipment. We therefore seek to explore a new purely visual attractor method that does not rely on specialized hardware and allows for walk-up interaction without prior calibration.

Driven by these considerations, we decided to thoroughly investigate flicker for guiding the observer through a dynamic narrative visualization. There are different visual features that can be oscillated to generate a flicker effect: geometric features (e.g., object shape or size) and color (e.g., hue and luminance). Since we were initially motivated by guidance through molecular processes (see Section 9), where preserving the geometry is crucial for an in-depth analysis of molecular structures and binding sites, we concentrated on color flicker. After informal pilot experiments, we could confirm that luminance flicker is a much stronger visual attractor than hue flicker. This is consistent with findings by Bailey *et al.* [2], and can be explained by the chromatic contrast blindness of humans in the peripheral vision [32]. We will therefore put the focus on luminance flicker in this paper.

4 COMPARATIVE PILOT EXPERIMENT

The goal of this initial pilot experiment was to verify that luminance flicker is indeed a strong visual attractor, as shown by prior work [34, 11, 16], and to investigate its effectiveness in a large, complex visualization. With this experiment, we aim to extend findings by Pinto *et al.* [34], who found that flickering targets can be easily spotted within uniformly moving distractors, by two aspects: first, an increased scene complexity by adding different colors and random object movements, and second, an extended visual field covered by the visualization. Our work also extends findings from experiments in desktop interface research (e.g., [4, 11, 16]) by comparing visual guidance techniques not only on a large display, but also in a complex, dynamic scene. For this purpose, we asked users to participate in a dual-task experiment in a scene with 1000 colored and randomly moving dots on a multi-monitor setup, comparing luminance flicker to two static visual guidance techniques.

The traditional approach to analyze visual attention is to employ eye tracking. However, the sole analysis of eye movements fails to explain cognitive processes of the observer while viewing the scene [5, 15], since eye movement patterns do not necessarily affect the user’s cognitive processing of the scene [7]. We therefore use other well-known metrics to quantify the amount of attention attraction, such as subjective reports and reaction times [5, 15] over eye tracking in all our experiments.

4.1 Participants

Ten participants (aged 21 to 50, 3 females) were recruited from a local university and a nearby company, including business assistants, software developers, and computer scientists.

4.2 Apparatus

We arranged three 27 inch monitors so that they subtended a visual angle of approximately 180° , as visualized in Fig. 1. We curved the monitors around the user so all pixels are equally usable, as recommended by Shupp *et al.* [38]. We used three identical BenQ BL2710 LCD monitors which were calibrated with an external monitor calibrator using a seven-color sensor. In sum, the three monitors provided 7680×1440 pixels. Users were not fixated into a chin rest and were instructed to behave naturally.

To compare attention attraction for conventional display sizes and large display setups covering also the peripheral visual field, we split the available display space into two separate regions. We used Sanders’ [36] separation of *head field* (i.e., where combined head-eye

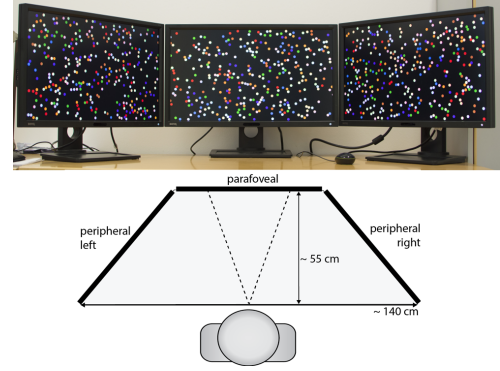


Fig. 1. Setup with three 27 inch monitors used in our experiments.

movement is required to acquire two subsequent targets) and *eye field* and *stationary field*, respectively (i.e., where subsequent targets can be spotted without head movements), to distinguish these two regions. Depending on the complexity of the scene, the *head field* starts at approximately 60° [36]. We therefore chose the two outer monitors – both starting at a subtended visual angle of $\sim 66^\circ$ – for displaying *peripheral* targets. To clearly distinguish peripheral from central targets, we defined a (*para*)*foveal* display region in the center of the central monitor, covering a subtended visual angle of $\sim 40^\circ$, as indicated by the dashed lines in Fig. 1.

During the experiment, users were exposed to a scene showing 1000 randomly moving and colliding dots with 20 pixels radius each. On the central monitor, one dot subtended a visual angle of $\sim 0.96^\circ$, which is about half of the foveal visual field with 2° . The number and size of dots was chosen to closely represent the appearance of well-known visual attention experiments (see [15] for an overview). To add complexity to such a scene, we added heterogeneity to two visual features: color and motion. The dots were colored by 12 different colors from a qualitative *ColorBrewer* [13] scheme on black background. Movement of dots was defined by random impulse forces applied routinely in random direction. The parameters of the random motion are the magnitude of the impulse and the friction coefficient. The parameters were chosen empirically so that dots would move between 21 to 170 pixels per second, which with our monitor resolution results in a speed the human visual system is most sensitive to [9]. The refresh rate of the visualization was synchronized with the monitor refresh rate of 60 Hz. Fig. 1(top) shows the scene on the three monitors.

4.3 Design

To prevent users from actively searching for a visual guidance signal, we employed a dual-task design. The primary task was to select as many red dots as possible on the central monitor. The secondary task was to press the space bar as quickly as possible when spotting a so-called “bonus-dot” appearing anywhere on the three monitors. The bonus dot was highlighted by one of three visual guidance techniques:

In the **flicker** technique, the bonus dot was signaled by a 10 Hz luminance flicker, covering the full luminance range from 0 to 100 in the CIE Lab space.

In the **spotlight** technique, presented by Khan *et al.* [22], the bonus dot was surrounded by a bright ring with a luminance value of 50, with twice the radius of the dot. In addition, the luminance of context dots was decreased by 75% (Fig. 2 left).

Finally, the **halo** technique used a ring identical to spotlight, but without context darkening (Fig. 2 right).

We used a 3×2 within-subjects design with the following factors:

Technique for visual guidance, as described above and **eccentricity** of the bonus dot on the screen (either (*para*)*foveal* or *peripheral*, as shown in Fig. 1).

Users had to perform ten repetitions, resulting in 60 trials in total. The trials were blocked per technique, and the order of the techniques was counter-balanced. The exact spatial location of the bonus

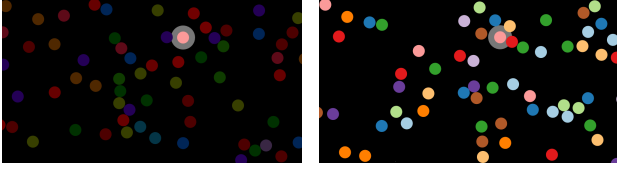


Fig. 2. Small sub-window with a spotlight (left) and halo (right).

dot within the (para)foveal or peripheral display regions, as well as the color of the dot, were picked randomly for each stimulus.

4.4 Procedure

For each trial, users had to press the Return-key to start a two-second countdown animation in the center of the screen. After this countdown, users could start with clicking red dots as their primary task. After a random period between two and six seconds, the bonus dot was highlighted. As soon as they pressed the space bar, or after a five-seconds timeout after the appearance of the highlight, the stimulus image was frozen, darkened, and all highlights were removed. Users then had to select the bonus dot with the mouse cursor, or press “n” in case they did not see it. Before the experiment, we demonstrated all three visual guidance techniques in their order of appearance and asked users to perform a training run. In the experiment, we logged *distance* (in pixels) of the user’s selection to the bonus dot (or -1, if the user pressed “n”) as a measure of correctness and *reaction time* between the highlight appearance and pressing space. In addition, we performed an audio-recorded post-experiment interview, where we asked people about perceived differences between the techniques.

4.5 Results and Discussion

The *success* of a trial was either 0, if the user pressed “n” or if the distance of the selection click to the actual bonus dot center was larger than 80 pixels (*i.e.*, 4 times the dot radius), or 1. The success rates were compared by a repeated measures logistic regression. Both, technique and eccentricity had a significant main effect (technique: $\chi^2(2) = 741.907, p < .001$; eccentricity: $\chi^2(1) = 901.518, p < .001$). Pairwise Bonferroni-corrected comparisons showed that halo led to a significantly lower success rate than spotlight and flicker, and that (para)foveal targets scored a higher success rate than peripheral ones (*cf.*, Fig. 3 left).

We conducted a mixed model analysis of reaction time of successful trials with technique and eccentricity as fixed factors and user as random factor. Both fixed factors are significant (technique: $F_{2,509.644} = 37.730, p < .001$; eccentricity: $F_{1,509.319} = 87.238, p < .001$), and there is also an interaction between them ($F_{2,509.438} = 13.771, p < .001$). Bonferroni-corrected post-hoc comparisons revealed that spotlight led to the fastest reaction times and halo to the slowest. (Para)foveal targets could be detected significantly faster than peripheral ones (*cf.*, Fig. 3 right).

In summary, the findings of our pilot experiment are:

First, targets in the (para)foveal region could be detected more reliably and faster than those in the periphery.

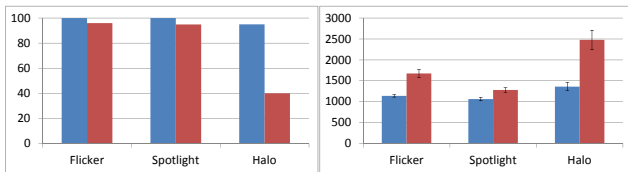


Fig. 3. Success rate in % for (para)foveal (blue) and peripheral (red) targets (left) and average reaction time in milliseconds with standard error bars (right).

Second, a simple halo around the target was more likely to be missed than a flickering target or a bright halo in a dark scene, and also led to the slowest reaction times. In the post-experiment interview, all users stated that halo was the hardest technique.

Third, flicker had an equal success rate as spotlight but led to a slower reaction time. Users explained that with spotlight “you just got a visual cue: now it happened!” (User 10). Subsequently, “you had to search around where the dot is exactly” (User 8), “...which was not very hard” (User 9). In contrast, with flicker “I don’t have to search, really, because I already know in which direction the blinking was” (User 7). However, a few users had the impression that spotlight caused a too obtrusive interruption (*e.g.*, User 5: “It is slightly irritating, because it covers the other things.”). While this darkening did not influence the performance in our simple experimental task, this strong interruption signal may be undesirable in many narrative visualizations, where users should be able to choose whether to follow the story, or to explore the scene without visual guidance.

Our findings suggest that luminance flicker can achieve a similar performance as the spotlight technique [22] in large, dynamic visualizations, even though the modifications to the scene are much more subtle. This extends findings from prior work by showing that flicker is also an effective attractor in large, dynamic scenes. But given that users rated flicker as disturbing in prior work [16], can we find a way to make it more appealing for narrative visualizations, where observers will be exposed to visual guidance for a longer time?

5 ATTRACTIVE FLICKER

The goal of Attractive Flicker is to exploit the effectiveness of luminance flicker for visual guidance, as demonstrated in the preceding pilot experiment, while decreasing the disturbance of the signal for a long-term exposure. The basic idea is to split the signal into two distinct stages, corresponding to the two visual attention stages by human observers described by Healey and Enns [15]: In the initial *orientation stage*, a short but sufficiently strong signal effectively guides the user’s attention to the focus element. In the subsequent *engagement phase*, the signal strength is smoothly reduced and remains at a constantly low level to minimize disturbance. What needs to be investigated is how long the initial orientation stage needs to be, as well as the flicker parameters for the two stages.

While the color of the focus element (given as L^* , a^* , and b^* $\in [0, 100]$) and the time t that passed since the element gained focus are defined by the narrative visualization, Attractive Flicker has two properties that define the signal’s perceived saliency, as well as its annoyance: amplitude and frequency. Since Attractive Flicker is split into two stages with different signal strengths, there are six parameters that define the overall signal: peak-to-peak amplitudes A_o and A_e (both $\in [0, 100]$) for the orientation and engagement stage, respectively, the periods T_o and T_e (where T is the reciprocal of the signal’s frequency f , given in milliseconds), and durations d_o of the orientation stage and d_t of a smooth transition period between orientation and engagement signal in milliseconds.

The luminance value $L^*(t)$ for a focus element subject to Attractive Flicker with a given $L^*a^*b^*$ color at time t is described by a wave

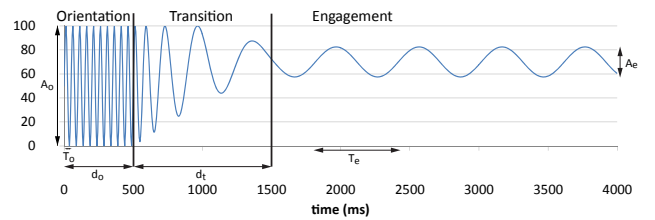


Fig. 4. Exemplary Attractive Flicker signal (four seconds) of a focus element of luminance $L^* = 70$ with parameters $A_o = 100$, $T_o = 50ms$, $A_e = 25$, $T_e = 600ms$, $d_o = 500ms$, and $d_t = 1000ms$.

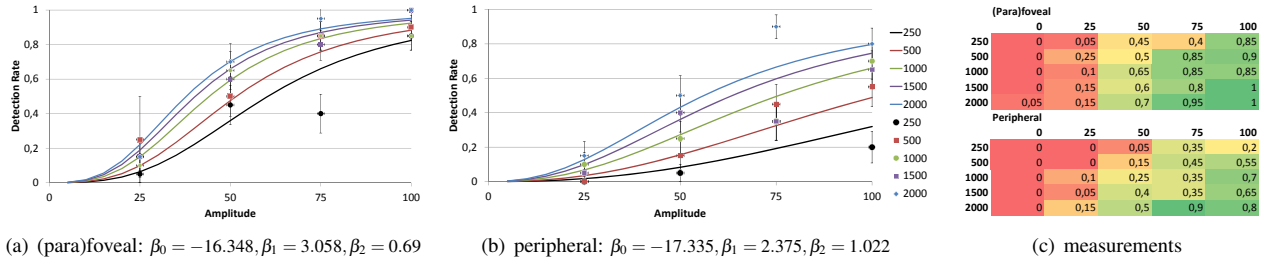


Fig. 5. Models predicting the detection rate for (a) (para)foveal and (b) peripheral targets, along with the estimated coefficients for Equation 3. Separate lines show the predicted detection rates per tested duration. Dots show the measured samples with standard error bars. (c) Average detection rates for (para)foveal (top) and peripheral targets (bottom), with durations as rows and amplitudes as columns.

function:

$$L^*(t) = L^{*f}(t) + \frac{A(t)}{2} \sin\left(2\pi \int_0^t \frac{1}{T(\tau)} d\tau\right), \quad (1)$$

where $A(t)$ is the peak-to-peak amplitude and $T(\tau)$ is the period at time $\tau \leq t$. In the transition period, the amplitude $A(t)$ and the period $T(t)$ with $d_o \leq t < d_o + d_t$ are linearly interpolated between A_o and A_e and T_o and T_e , respectively. The center of the wave amplitude L^{*f} at a given time t is defined as:

$$L^{*f}(t) = \begin{cases} \min\{L^* + \frac{A(t)}{2}, 100\} - \frac{A(t)}{2}, & \text{if } L^* \geq 50 \\ \max\{L^* - \frac{A(t)}{2}, 0\} + \frac{A(t)}{2}, & \text{otherwise,} \end{cases} \quad (2)$$

which assures that all colors have the same flicker amplitude, irrespective of the element's luminance. Fig. 4 illustrates an exemplary Attractive Flicker signal. Candidates for the signal parameters A_o , T_o , A_e , T_e , d_o , and d_t were determined in two perceptual studies, presented in the following sections.

6 PERCEPTUAL STUDY 1: ATTENTION ATTRACTION

The goal of the first perceptual study was to empirically determine the minimally required flicker amplitude A_o and duration d_o to reliably guide the user's attention to both, (para)foveal and peripheral display regions, in the orientation stage. For this purpose, users were asked to participate in a visual search task for a flickering target on a large display containing 1000 randomly moving dots. By systematically varying the flicker amplitude of the target and the stimulus duration, our aim was to find candidates for these parameters so that the detection rate is above a pre-defined threshold of 90%. The hardware setup and scene settings were identical as in the comparative pilot study (Section 4).

6.1 Participants

Ten participants (aged 25 to 37, one female) were recruited from a local university, including research scientists in computer science and technical assistants.

6.2 Design

We employed a $2 \times 5 \times 5$ within-subjects design with the following factors:

Eccentricity of the target (either (para)foveal or peripheral, as described in Section 4),

amplitude A of the luminance modulation (as described in Equation 1), defined by the peak-to-peak modulation range between 0 and 100 in CIELab space (0, 25, 50, 75, 100, where 0-values represented target-absent trials), and

duration d of the stimulus presentation in milliseconds (250, 500, 1000, 1500, 2000).

The duration values were chosen so that the shortest duration was slightly longer than a minimum dwell period (≥ 200 ms) and that the maximum tested duration was clearly longer than saccadic movements

followed by head movements (≥ 500 ms) [43]. Amplitude values covered a range from borderline visibility determined in an informal pilot test ($A = 25$) to the full possible luminance amplitude ($A = 100$).

Users had to perform two repetitions, resulting in 100 trials in total, where the order of appearance of the 50 different stimuli was randomized. In each stimulus, there was exactly one blinking dot — or none for target-absent trials. Flicker frequency was fixed to 20 Hz, since informal pilot tests indicated that slow blinking was generally much harder to detect (see also [4, 11]).

6.3 Procedure

Before each trial, the user pressed the Return-key to start a two-second countdown animation in the center of the screen. Users were instructed to fixate the countdown animation until the stimulus would appear. Each stimulus was shown according to the duration factor, i.e., between 250 and 2000 milliseconds. After this duration, the image was frozen and darkened. Users then had to mark the approximate area where they perceived some flicker before the stimulus end with the mouse cursor, or press "n" in case they did not perceive any flicker. We logged *detection* (0 or 1) and, if the flickering dot was detected, *accuracy* by the distance of the user's indicated flicker location and the actual position of the target in pixels.

6.4 Results

We first performed a Wilcoxon Signed-Rank test to compare average detection rates between (para)foveal and peripheral display regions. We found a significant difference for eccentricity ($Z = -2.803, p = .005$), with the average detection rate being 43% lower in peripheral display regions than in the (para)foveal region. All subsequent analyses were performed separately for (para)foveal and peripheral regions.

To analyze the effect of flicker amplitude A and signal duration d on the detection rate, we performed a binary logistic regression with logarithmic A and d as predictors:

$$dr(A, d) = \frac{\exp(\beta_0 + \beta_1 \ln(A) + \beta_2 \ln(d))}{1 + \exp(\beta_0 + \beta_1 \ln(A) + \beta_2 \ln(d))} = \frac{1}{1 + e^{-\beta_0 A - \beta_1 d - \beta_2}}. \quad (3)$$

The estimated coefficients and the resulting predictions are shown in Fig. 5. The Nagelkerke- R^2 goodness-of-fit measures indicate moderate (.472 for the (para)foveal model) and modest (.35 for the peripheral model) fits of model to data, respectively.

As indicated in Fig. 5, for the (para)foveal display regions, our model predicts a detection rate of $\geq 90\%$ with flicker duration $d \geq 650$ and amplitude $A = 100$. In peripheral display regions, the maximum predicted detection rate for the maximum flicker duration $d = 2000$ and amplitude $A = 100$ is slightly below 80%.

Accuracy of target location estimation was fairly high, with 74.28% of all detected targets being directly selected. For those detected dots that were not accurately selected, the average distance to the target was 146.3 pixels. For (para)foveal display regions, users could accurately determine the location of all detected targets for all trials with duration $d \geq 1000$, independently of the amplitude. For peripheral targets,

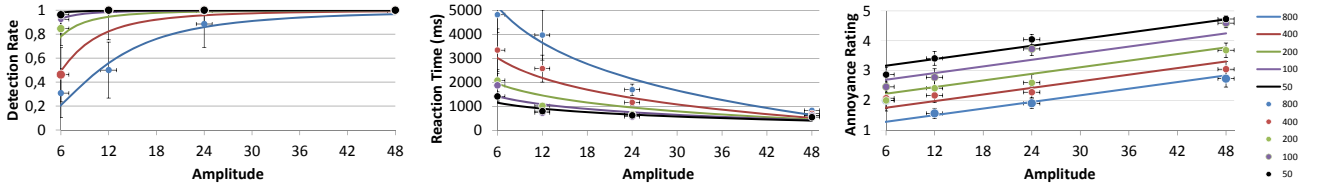


Fig. 6. Models predicting the (left) detection rate, (middle) reaction time, and (right) annoyance. Separate lines show the predicted models per tested periods. Dots show the measured samples with standard error bars.

only 67.8% of targets could be accurately spotted, while the average distance to the remaining targets was 180.7 pixels.

7 PERCEPTUAL STUDY 2: MINIMALLY DISTURBING FLICKER

The goal of the second perceptual study was to investigate the trade-off between flicker effectiveness and subjective annoyance in the engagement stage by investigating two parameters: flicker amplitude A_e and period T_e . In a visual search task, we aimed to find promising amplitude and period ranges by evaluating a cost-benefit function which trades off a high detection rate (benefit) against increased reaction time and annoyance (cost).

7.1 Participants

Eleven participants (aged 23 to 33, four females) were recruited from a local university. All participants were master students or post-graduates of computer science.

7.2 Apparatus

The stimuli were presented in a 300×300 pixels window on the central monitor of Fig. 1. The size of this window was chosen so that the search radius would correspond to the average distance to targets that could not be uniquely identified in the first perceptual study (~ 150 pixels). The window contained 12 randomly moving and colliding dots of same size, colors, and movement speed as in the previous studies.

7.3 Design

We employed a 5×5 within-subjects design with the following factors: **Amplitude** A of the luminance modulation, defined by the peak-to-peak modulation range between 0 and 100 in CIELab space (0, 6, 12, 24, 48, where 0-values represented target-absent trials), and **period** T of the flicker signal in milliseconds (800, 400, 200, 100, 50). The period value 50 corresponds to the flicker frequency $f = 20$ Hz used in the first perceptual study.

Users had to perform two repetitions, resulting in 50 trials in total, where the presentation order of the 25 stimuli was randomized. Color and spatial location of the target dot within the window was picked randomly.

7.4 Procedure

As in the previous perceptual study, the users could initiate a new trial by hitting the Return-key, followed by a two-second countdown visualization. Then, the stimulus was shown for a maximum duration of 10 seconds. Users were asked to hit the space key as soon as they spotted a flickering dot. After hitting the space key or reaching the time-out, the stimulus was frozen and darkened, and users had to pick the previously flickering dot with the mouse cursor, or press “n”, if they did not see any flickering. If a dot was selected, the stimulus was restarted and users were asked to watch the scene until they felt confident to rate the annoyance of the signal with a value between 1 (not annoying at all) and 5 (extremely annoying) on the numpad. We instructed the users that annoyance should be rated with the assumption that the signal would be shown to them for a longer time, following a methodology by Bartram *et al.* [4]. For each trial, we logged target *detection* (0 or 1) and, if the flickering dot was detected, *reaction time*, the subjective *annoyance* rating, and *correctness* by evaluating the distance of

the picked dot location to the actual target dot center, where all picked locations ≤ 25 pixels from the target’s center were counted as correct.

7.5 Results

To find a good trade-off between detection, annoyance, and reaction time, we defined a cost-benefit function. We optimized the cost-benefit subject to the desired constraint that the detection rate is ≥ 0.9 , according to the reliability threshold we used in the first perceptual study:

$$cb(A, T) = \begin{cases} \frac{dr(A, T)}{an(A, T) \cdot rt(A, T)}, & \text{if } dr(A, T) \geq 0.9, \\ 0, & \text{otherwise,} \end{cases}$$

where $dr(A, T)$, $rt(A, T)$, and $an(A, T)$ are predicting the detection rate, reaction time, and annoyance, respectively, as a function of A and T , as described below:

The detection rate $dr(A, T)$ is described by a binary logistic regression, as given in Equation 3. The Nagelkerke- R^2 goodness-of-fit measure reveals that the model fits moderately well to the data (.51). The coefficients ($\beta_0 = 7.192, \beta_1 = 2.268, \beta_2 = -1.884$) indicate that amplitude is a slightly more reliable predictor of the detection rate than the signal period. Our model predicts that for all our tested amplitudes $A \geq 6$, we could achieve 90% detection rate for periods $T \leq 100$. For amplitudes $A \geq 30$, all periods lead to a detection rate above 90%. Fig. 6 (left) visualizes the detection rate model, as well as the experiment samples.

Of all detected targets, 92.2% were correct. The analyses of reaction times and annoyance were performed only for the correctly detected targets:

To model the reaction time, the best-fitting model ($R^2 = .426$) was achieved by a regression, with the following estimated coefficients:

$$rt(A, T) = 1317.423 - 238.143 \cdot \ln(A) + 9.682 \cdot T - 2.42 \cdot \ln(A) \cdot T.$$

As visualized in Fig. 6 (middle), the reaction time decreases with logarithmically increasing amplitude and linearly decreasing period. Also, there is an interaction between amplitude and period. While reaction times strongly deviate for small amplitudes depending on the frequency, they all converge to around 500 milliseconds for the highest measured amplitude $A = 48$.

Annoyance was modeled by a regression with the following best-fitting model ($R^2 = .359$) and estimated coefficients:

$$an(A, T) = 5.598 + 0.037 \cdot A - 0.679 \cdot \ln(T).$$

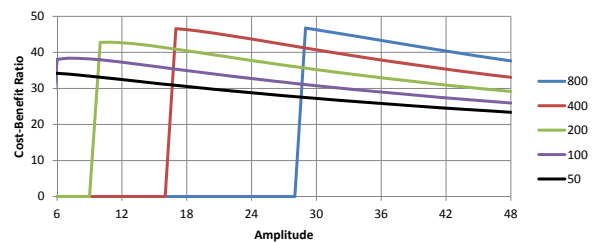


Fig. 7. Cost-benefit function for flicker amplitudes and tested periods (separate lines); higher values are better.

The model predicts that annoyance increases with linearly increasing amplitude and logarithmically decreasing period (*cf.*, Fig. 6 right).

From observing Fig. 7, we can determine that the maximum of the cost-benefit function is reached for periods between $T = 400$ and $T = 800$ with amplitudes between $A = 17$ and $A = 29$.

8 DISCUSSION OF PERCEPTUAL STUDY RESULTS

Like the initial pilot experiment, our first perceptual study showed that the detection of a flickering target is more reliable in (para)foveal display regions as compared to peripheral regions. But even in the (para)foveal region, only scene elements flickering with the highest possible luminance amplitude of $A = 100$ could reliably attract the visual attention of the observer within 500 milliseconds in our complex scene. It seems that with lower amplitudes, the popout effect of flickering dots in our scene was too weak to be perceived pre-attentively. In peripheral display regions, we could not reach our desired 90% detection rate within our tested stimulus durations of $d \leq 2000$ milliseconds. This implies that luminance flicker alone may not be sufficient for reliably attracting the attention to peripheral display regions within the first two seconds of stimulus onset. In the future, it will therefore be necessary to investigate different flicker features, or combinations of features, to guide the attention more effectively to these peripheral areas.

If detected, users could point at target locations fairly accurately, even after short flicker durations. Unsurprisingly, the accuracy increased with the stimulus duration. On the peripheral monitors, users often could not indicate the exact target, but were able to identify a close area around it. This is sufficient, since our two-stage signal should take care that the user’s gaze will be guided to the correct target within this small search window after the initial orientation stage.

For the small search window tested in the second experiment, we found that a much smaller amplitude and lower frequency is sufficient to be able to reliably detect a target. However, with increasing detection rate and shorter reaction times, subjective annoyance is also increasing. In other words: the more effective a signal, the more disturbing it is perceived. This is consistent with prior experiments [4, 11, 16]. Our results indicate that a good trade-off between detection reliability and annoyance can be found for the low frequency and medium amplitude ranges tested in our experiment.

Our results provide some guidance for the choice of suitable amplitude and frequency candidates for both flicker stages, as well as the minimum duration of the orientation stage. However, there are several factors in a visualization that will influence these candidate parameters:

First, while we varied the dots’ colors and motion directions in our experiments, they all had the same size, each stimulus contained the same number and density of dots, and the distance of the user to the three displays was quite homogeneous. It can be expected that the size of the target has a strong influence on the minimally required amplitude and frequency parameters for reliable detection. Similarly, the much higher detection rates in the second perceptual study suggest that the size of the search window and the number of context elements, respectively, play a major role in the perceived flicker strength. Systematically exploring how the interplay between focus and context size influences the flicker parameters will be an important future work.

Second, the dots in our scene were moving in random directions, but with homogeneous velocities. Variations of movements speeds are likely to also influence the popout effect of flickering targets in the orientation stage.

Third, due to the task descriptions and the high number of repetitions, users’ top-down attention was tuned to detect flicker, which is an inherent property of classic target detection experiments (*cf.*, [15]). However, when exploring a guided narrative visualization, we assume that users will also learn to expect a certain repeating popup feature.

Forth, the relative annoyance ratings of the second perceptual study only relate to the signal itself, not to the signal in context of a real task. In real scenarios, the annoyance caused by an attractor also depends on its perceived usefulness. Therefore, researchers have recommended to adjust the attractor strength to its interruption importance [11]. In the

case of Attractive Flicker, the duration of the orientation stage, as well as amplitude and frequency of the engagement stage could be adjusted to the predicted degree of interest of the target.

Finally, we experimented with amplitude and frequency for luminance flicker. Even though we expect similar results for parameters of alternative flicker features, more empirical evidence is required in this regard.

9 EXPERIMENT: MOLECULAR STORIES

We demonstrate the effectiveness of Attractive Flicker in an exemplary showcase of molecular visualization. According to Jenkison and McGill [21], a more complex representation of a molecular scene — showcasing a lot of details like random motion or molecular crowding — offers a better understanding of the visualized process than a more simplistic abstraction. This strongly motivates the use of guided narrative storytelling in molecular dynamic simulations with a very detailed molecular scenery, while the observer is visually guided through the narrative of a biochemical process.

Complex systems like these are often simulated by agent-based models. In such systems, events take place punctually at distinct times and locations. To be able to grasp the visualization’s narrative, the user needs visual guidance to track the ongoing process in focus, but also requires basic understanding of the embedding surroundings.

As an example scenario of the molecular machinery, we simulate and visualize a simplified version of the *nicotinamide adenine dinucleotide* (NAD) synthesis process in the *salvage pathway*. NAD is a molecule present in the cells of all living organisms, mostly used in electron transfer reactions.

To assess the effectiveness of Attractive Flicker in such a scenario, we compare it to two baseline conditions: to the unmodified visualization and a radical context masking, where only the elements in focus are being displayed. We have two hypotheses: First, we hypothesize that using Attractive Flicker, users will be able to follow the visualization’s narrative as well as if only the focus elements were displayed without any context. Second, we hypothesize that Attractive Flicker will be minimally distracting, so perception of contextual information will be as good as with the unmodified scene without any visual guidance.

9.1 Participants

Twelve paid participants (aged 23 to 32, four females), including computer scientists, students, and teachers of different disciplines, participated in the study.

9.2 Apparatus

The experiment was conducted on the three-monitor setup shown in Fig. 1. In a 7680×1440 pixels full-screen window, we displayed a 3D molecular scene containing four different types of enzymes and ten different types of metabolites of the NAD synthesis process. Since metabolites were quite small and had similar shapes, we used an adapted qualitative color-coding from *ColorBrewer* [13] to distinguish them. Enzymes were rendered in grey. We set the quantities of all but two metabolites to 100 elements per scene. One metabolite quantity was set to 800, one reaction product was set to 0 initially. In total, the scene contained ~ 2000 elements (*cf.*, Fig. 8).

In each trial, the simulation triggered a chain of three reactions out of the salvage pathway, which was replayed four times. The first reaction started approximately five seconds after the trial initiated. Between each reaction, there was a break of 0 to 15 seconds. Since metabolites and enzymes were subject to random forces, the duration of the reactions varied. In total, each trial scene was displayed for four to five minutes.

To ensure the visibility of reactions, the depth of the volume was limited to avoid occlusions of enzymes. Whenever an enzyme was selected as reaction catalyst, we rotated the enzyme so its binding site was facing towards the user and the docking metabolites were clearly visible.

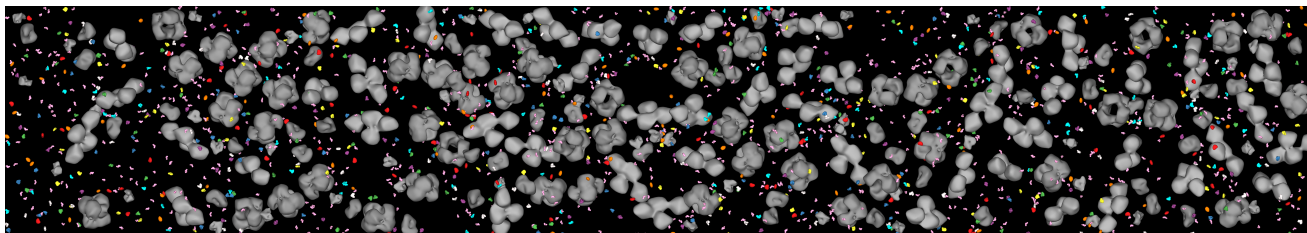


Fig. 8. Screenshot of the dynamic visualization showing the molecular machinery with four different enzymes and ten different metabolites.

9.3 Design

Since we were interested in the users’ insights generated from the displayed visualization, we evaluated this experiment by means of content questionnaires. We designed two classes of questions: The *focus questions* targeted towards the chain of reactions. For all three reactions in each trial, users had to identify the metabolites acting as substrates and the resulting products. In addition, they had to note the order of the three reactions. The *context questions* addressed overall scene properties, irrespective of the reactions taking place. For each trial, users had to find the metabolite with most and least instances in the scene, respectively. Since the molecules were presented purely visually without any textual labels, users had to select the correct metabolites by picking their color from a list of all 10 metabolite colors used in the scene, printed on the questionnaires.

We employed a within-subjects design with three visual guidance conditions:

In the **Normal** condition, we showed the unmodified scene without any visual hints supporting the detection of ongoing reactions.

In the **Attractive Flicker** condition, the material color of the reaction partners (*i.e.*, the enzyme and one or two substrates) was oscillated using our Attractive Flicker technique. According to the findings in our perceptual studies, we displayed the orientation phase with full amplitude $A_o = 100$ and a short period $T_o = 50$ for 500 milliseconds. For the subsequent engagement period, we used an amplitude of $A_e = 25$ and a period of $T_e = 600$. The two signals were connected by a transition period of 1500 milliseconds.

In the **No Context** condition, we displayed only the reaction partners and completely masked the remaining scene. This can be compared to an extreme *spotlight* effect (*cf.*, Section 4). Mind that we did not issue the context questions for the *No Context* condition.

The presentation order of the three conditions was counter-balanced. In addition, we constructed two scenes with slight variations of the salvage pathway, and changed the color assignments to the metabolites for each scene. The assignment of these three resulting scenes to the three visual guidance conditions was also counter-balanced. While the color-coding and quantities of the metabolites, as well as the chain of reactions, were pre-defined for each scene, the spatial locations of the reactions on the display were picked randomly.

9.4 Procedure

Before starting the first trial, we showed the user a short demonstration scene in the *No Context* condition to explain the visual appearance of the individual reactants, as well as how a reaction would look like. Before each trial, users were issued a short textual description of the current visual guidance condition. We avoided any mentioning of the term “flicker” in these scene descriptions, and consistently used the term “highlighting” instead, to avoid tuning the top-down attention of participants to flickering elements. Then, users could first conduct a test run with a training data set. After the experiment, we asked users to rate the task difficulty for both, focus and context questions, for each visual guidance technique on a five-point Likert scale. Finally, we gathered qualitative feedback in an audio-recorded semi-structured interview by asking for more details about the subjective questionnaire ratings and whether users had any additional comments.

9.5 Results

We will report the correctness of the answers for focus questions, as well as their perceived difficulty, for all three guidance conditions. Results for the context questions and their perceived difficulty will only be compared between *Attractive Flicker* and the *Normal* scene.

9.5.1 Focus Questions

To determine the correctness of the noted reactions, we counted the number of correctly identified reactants, where a maximum of four points could be scored for each reaction. A correct sequence was scored when two subsequent reactions were reported in their correct order, which leads to a maximum value of two correct sequences for the three reactions per trial.

We performed a one-way ANOVA with Bonferroni-corrected post-hoc comparisons on correctness of reaction partners across the three visual guidance conditions. We found a significant difference between the three conditions ($F_{1,113,12,245} = 580.833, p < .001, \eta^2 = .981$). Post-hoc comparisons revealed that both, *No Context* and *Attractive Flicker* (both 99.3% correctness), yielded significantly better results than *Normal* (6.9% correctness). Similarly, we found a significant difference for the correctness of reaction sequences ($F_{2,22} = 209.0, p < .001, \eta^2 = .95$). *Attractive Flicker* (87.5% correctness) led to a significantly better result than the *Normal* scene. In the *No Context* condition, all sequences were reported correctly, while in the *Normal* scene, no correct sequence was found.

The post-experiment questionnaire was evaluated using a Friedman test with Bonferroni-corrected Wilcoxon Signed-Rank post-hoc comparisons. The ease of finding reactions was rated significantly differently for the three conditions ($\chi^2(2) = 21.565, p < .001$): The average score for *No Context* of 4.92 was better than for *Attractive Flicker* with an average score of 4.08 ($Z = -2.486, p = .013$). *Normal* reached the lowest average score with 1.17, which is significantly different from both, *No Context* ($Z = -3.217, p = .001$) and *Attractive Flicker* ($Z = -3.108, p = .002$).

In the post-experiment interview, all users stated that finding reactions in the *Normal* scene was “*not or close to impossible*”. They also reported several reasons why finding and tracking reactions was slightly harder using *Attractive Flicker* than *No Context*: distraction and occlusions caused by surrounding molecules, too subtle flickering of the small metabolites, and difficulties when re-acquiring the reaction partners after switching the attention to the questionnaire. Therefore, three users suggested to add pulsating halos around the focus metabolites to put more visual emphasis especially on small targets.

These results suggest that our first hypothesis is supported: Users could answer questions about the story in focus as accurately as with a scene completely masking the context. However, it was less demanding for the users to follow the story with complete context masking.

9.5.2 Context Questions

For the quantity estimations, users could score one point for each trial for both, the highest and lowest quantity estimation. We compared the estimations using a Wilcoxon Signed-Rank test. We neither found a significant difference for the highest quantity estimations ($Z = -1.000, p = .317$ with 100% correctness for *Normal* and 91.7%

correctness for *Attractive Flicker*), nor for the lowest quantity estimations ($Z = -.577, p = .564$, with 91.6% correctness for *Normal* and 83.3% correctness for *Attractive Flicker*).

This is also reflected in the post-experiment questionnaires. A Wilcoxon Signed-Rank test revealed that users rated the ease of finding the highest and lowest concentrations equally easy ($Z = -.378, p = 0.705$), with an average score of 4.25 for the *Normal* scene and 4.17 for the *Attractive Flicker* scene. In the interview, eight users reported that the flickering did not distract them when estimating the quantities (for instance, User 11: “*The flickering, I can simply ignore.*”). No user reported that the flickering reactions were disturbing when observing the scene context.

These findings confirm our second hypothesis: Attractive Flicker adds negligible distraction when concentrating on the visualization’s context.

10 DISCUSSION

With these experimental results, we will first discuss the design of Attractive Flicker, followed by an analysis of additional use cases and open research questions.

10.1 Two-Stage Signal

We proposed a two-stage signal to decrease the strong attention guidance effect of flicker to a persistent signal with significantly lower subjective disturbance. We showcased this approach with luminance flicker to guide the user through a dynamic narrative visualization on a large display. However, this proposed two-stage design is not limited to luminance flicker. The same principle can be applied for all different sort of visual attractors that increase the visual emphasis of the focus object. For instance, when applying the two-stage concept to halos around focus elements, a large and clearly visible bright highlight around the focus object will dynamically shrink to a small, subtle halo in the engagement stage. Similarly as demonstrated for luminance flicker in this paper, the required parameters for the orientation and engagement stages (e.g., radius and brightness difference for the two-stage halo), need to be explored empirically.

In our first perceptual study, we explored how long the orientation stage needs to last for the flicker attractor to be reliably perceivable and applied our findings in a case study of molecular stories. Although users could easily follow the story when guided by Attractive Flicker, some users reported that they were afraid to miss the initial signal in some situations. It seems that, even though the orientation stage signal can be perceived reliably, users sometimes would like to have a longer orientation stage in realistic scenarios. Another potential way to set a useful orientation stage duration could be a combination of Attractive Flicker with an eye tracking approach (as proposed by Bailey *et al.* [2]) or even a coarse head tracking technology using an off-the-shelf webcam. This way, the transition phase would be initiated only after the user directs the gaze towards the highlighted target.

10.2 Narrative Visualizations and Other Use Cases

The field of narrative visualization has generated increased interest in recent years [46, 37, 27]. According to Segel and Heer [37], *highlighting* to guide the user’s attention to elements of interest is one of three *visual narrative tactics*, besides providing the user with an overview (*visual structuring*) and *transition guidance* to change the visualization viewpoint or representation without disorienting the user. These tactics can co-exist in parallel. For instance, animated transitions can resolve occlusions, while visual highlighting guides the user’s gaze to the most relevant regions while preserving the context. The larger the scene, the more important highlighting is to guide the user’s attention through the narrative.

We demonstrated the usefulness of Attractive Flicker as highlight technique for large, dynamic scenes. In our use case, observers could as accurately follow the narrative of a story as if only the story without any context was being displayed. Notably, our users consistently reported that Attractive Flicker was easy to ignore when observing the context of the scene. Since only for the initial orientation stage we chose amplitude and frequency parameters high enough to generate

a popout effect on the large display, the potential source of distraction was limited to less than a second. However, when guiding a user through a smaller or static visualization, static attractors (like, for instance, a halo as in the comparative pilot experiment) are probably sufficiently distinct to generate a popout effect and will therefore cause a lower annoyance. On the other hand, when dealing with volume data, simply flickering a region’s luminance is not possible to make it stand out, due to potential occlusions. In such a scenario, a visual guidance technique can be implemented by utilizing and extending the concept of temporal transfer functions as proposed in the scientific visualization literature [28].

Narrative visualization is not only used to passively communicate a pre-defined story, as in our use case, but it is also an important tool to guide interactive exploration of complex information. For instance, when constantly analyzing a large amount of real-time data, such as during emergency management or stock broking, Attractive Flicker can help to make new and important data visually stand out in a minimally obtrusive way. Similarly, Attractive Flicker could serve as indicator of visual changes caused by user’s brushing actions in large-scale multiple coordinated views. For this application area, it will be necessary to explore the effectiveness of multiple synchronous Attractive Flicker targets, potentially with different strengths matched to the predicted degrees of interest of the elements.

Outside of the visualization field, educational animations, for instance, are in general controversial since learners have difficulties drawing their attention to semantically relevant regions without visual guidance [26, 24]. Traditionally, attention guidance also plays an importance role in desktop interfaces to effectively deliver notifications (cf., [4, 11, 16]) and advertisement. These application areas could also benefit from a technique that initially guides the attention for a short period of time, before decaying into a less distracting persistent state.

11 CONCLUSION

We presented Attractive Flicker — a visual attention guidance technique for large, complex visualizations based on luminance oscillation of focus elements. The novelty of our technique is that the flicker signal is split into two stages, inspired by perceptual attention stages of human observers. We could empirically determine that the initial orientation stage does not need to be longer than 500 milliseconds — i.e., below the average duration of a coordinated eye-head movement [43] — but very salient to reliably attract the user’s attention. We also found that small targets in the periphery cannot fully reliably attract the user’s attention with luminance flickering alone. For the subsequent engagement stage, we observed that an amplitude covering one quarter of the luminance range and a frequency below 2 Hz represents a good trade-off between signal effectiveness and subjective annoyance. In a showcase scenario of molecular interactions on a large display, we could verify that Attractive Flicker indeed adds minimal distraction, but is still a highly effective technique to guide the user through a narrative visualization.

In the future, we want to systematically investigate more scene parameters potentially influencing the effectiveness of Attractive Flicker, such as focus and context size and heterogeneity of scene motion or object shapes. In addition, we plan to extend our proposed two-stage model to other classes of visual guidance techniques, and explore combinations of flickering and static techniques.

ACKNOWLEDGMENTS

This research has been financed by the Vienna Science and Technology Fund (WWTF) through project VRG11-010, and additional support has been provided by the EC Marie Curie Career Integration Grant through project PCIG13-GA-2013-618680. We thank all our study participants for their support.

REFERENCES

- [1] Gapminder: Unveiling the beauty of statistics for a fact based world view. <http://www.gapminder.org/>, 2014.
- [2] R. Bailey, A. McNamara, N. Sudarsanam, and C. Grimm. Subtle gaze direction. *ACM Trans. Graph.*, 28(4):1–22, Sept. 2009.

- [3] L. Bartram and C. Ware. Filtering and brushing with motion. *Information Visualization*, 1(1):66–79, Mar. 2002.
- [4] L. Bartram, C. Ware, and T. Calvert. Moticons:: detection, distraction and task. *International Journal of Human-Computer Studies*, 58(5):515–545, May 2003.
- [5] A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, 2013.
- [6] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella. Directing gaze in 3d models with stylized focus. In *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, EGSR’06, pages 377–387. Eurographics Association, 2006.
- [7] B. B. de Koning, H. K. Tabbers, R. M. Rikers, and F. Paas. Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20(2):111–122, Apr. 2010.
- [8] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for focus+context visualization of complex simulation data. In *Proceedings of the symposium on Data visualisation 2003*, VISSYM ’03, page 239248, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.
- [9] E. N. Dzhabarov, R. Sekuler, and J. Allik. Detection of changes in speed and direction of motion: Reaction time analysis. *Perception & Psychophysics*, 54(6):733–750, Nov. 1993.
- [10] G. W. Furnas. Generalized fisheye views. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI ’86, pages 16–23. ACM, 1986.
- [11] J. Gluck, A. Bunt, and J. McGrenere. Matching attentional draw with utility in interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’07, pages 41–50. ACM, 2007.
- [12] C. Gutwin and C. Fedak. A comparison of fisheye lenses for interactive layout tasks. In *Proceedings of the 2004 Graphics Interface Conference*, GI ’04, pages 213–220. Canadian Human-Computer Communications Society, 2004.
- [13] M. Harrower and C. A. Brewer. ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, June 2003.
- [14] H. Hauser. Generalizing focus+context visualization. In *Scientific Visualization: The Visual Extraction of Knowledge from Data*, Mathematics and Visualization, pages 305–327. Springer Berlin Heidelberg, 2006.
- [15] C. G. Healey and J. Enns. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1170–1188, 2012.
- [16] R. Hoffmann, P. Baudisch, and D. S. Weld. Evaluating visual cues for window switching on large screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’08, pages 929–938. ACM, 2008.
- [17] D. Huber and C. G. Healey. Visualizing data with motion. In *IEEE Visualization, 2005. VIS 05*, pages 527–534, 2005.
- [18] L. Itti. Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12(6):1093–1123, 2005.
- [19] L. Itti, N. Dhavale, and F. Pighin. Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Optical Science and Technology, SPIE’s 48th Annual Meeting*, volume 5200, pages 64–78. International Society for Optics and Photonics, 2004.
- [20] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.
- [21] J. Jenkinson and G. McGill. Visualizing protein interactions and dynamics: Evolving a visual language for molecular animation. *CBE-Life Sciences Education*, 11(1):103–110, Mar. 2012.
- [22] A. Khan, J. Matejka, G. Fitzmaurice, and G. Kurtenbach. Spotlight: Directing users’ attention on large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’05, pages 791–798. ACM, 2005.
- [23] Y. Kim and A. Varshney. Saliency-guided enhancement for volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):925–932, Sept. 2006.
- [24] B. B. de Koning, H. K. Tabbers, R. M. J. P. Rikers, and F. Paas. Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21(2):113–140, June 2009.
- [25] R. Kosara, S. Miksch, and H. Hauser. Focus+context taken literally. *IEEE Computer Graphics and Applications*, 22(1):22–29, Jan. 2002.
- [26] R. K. Lowe. Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13(2):157–176, Apr. 2003.
- [27] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, Jan. 2012.
- [28] R. Maciejewski, I. Woo, W. Chen, and D. Ebert. Structuring feature space: A non-parametric method for volumetric transfer function generation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1473–1480, Nov. 2009.
- [29] A. McNamara, T. Booth, S. Sridharan, S. Caffey, C. Grimm, and R. Bailey. Directing gaze in narrative art. In *Proceedings of the ACM Symposium on Applied Perception*, SAP ’12, pages 63–70. ACM, 2012.
- [30] R. Minghim and A. R. Forrest. An illustrated analysis of sonification for scientific visualisation. In *Proceedings of the 6th Conference on Visualization ’95*, pages 110–117. IEEE Computer Society, 1995.
- [31] P. K. Mital, T. J. Smith, R. L. Hill, and J. M. Henderson. Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, 3(1):5–24, Mar. 2011.
- [32] J. Moreland and A. Cruz. Colour perception with the peripheral retina. *Optica Acta: International Journal of Optics*, 6(2):117–151, 1959.
- [33] H. E. Petersen and D. J. Dugas. The relative importance of contrast and motion in visual detection. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 14(3):207–216, June 1972.
- [34] Y. Pinto, C. N. L. Olivers, and J. Theeuwes. Selecting from dynamic environments: Attention distinguishes between blinking and moving. *Perception & Psychophysics*, 70(1):166–178, Jan. 2008.
- [35] R. Rosenholtz. A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39(19):3157–3163, Oct. 1999.
- [36] A. F. Sanders. Some aspects of the selective process in the functional visual field. *Ergonomics*, 13(1):101–117, 1970.
- [37] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.
- [38] L. Shupp, R. Ball, B. Yost, J. Booker, and C. North. Evaluation of viewport size and curvature of large, high-resolution displays. In *Proceedings of Graphics Interface 2006*, pages 123–130. Canadian Information Processing Society, 2006.
- [39] M. Steinberger, M. Waldner, A. Lex, M. Streit, and D. Schmalstieg. Context-preserving visual links. *IEEE TVCG 2011*, 17(12):2249–2258, 2011.
- [40] V. Sundareswaran, K. Wang, S. Chen, R. Behringer, J. McGee, C. Tam, and P. Zahorik. 3d audio augmented reality: Implementation and experiments. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 296–297. IEEE Computer Society, 2003.
- [41] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, Jan. 1980.
- [42] E. E. Veas, E. Mendez, S. K. Feiner, and D. Schmalstieg. Directing attention and influencing memory with visual saliency modulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’11, pages 1471–1480. ACM, 2011.
- [43] C. Ware. *Information Visualization: Perception for Design*. Elsevier, May 2012.
- [44] C. Ware and R. Bobrow. Motion to support rapid interactive queries on nodelink diagrams. *ACM Trans. Appl. Percept.*, 1(1):318, July 2004.
- [45] C. Ware and R. Bobrow. Motion coding for pattern detection. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, APGV ’06, pages 107–110. ACM, 2006.
- [46] M. Wohlfart and H. Hauser. Story telling for presentation in volume visualization. In *Proceedings of the 9th Joint Eurographics / IEEE VGTC Conference on Visualization*, EUROVIS’07, page 9198, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.
- [47] J. M. Wolfe, K. R. Cave, and S. L. Franzel. Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3):419–433, 1989.
- [48] S. Zhai, J. Wright, T. Selker, and S.-A. Keln. Graphical means of directing users attention in the visual interface. In *Human-Computer Interaction INTERACT 97*, IFIP The International Federation for Information Processing, pages 59–66. Springer US, Jan. 1997.

2.5 A Perceptual-Statistics Shading Model

Veronika Šoltészová, Çağatay Turkay, Mark C. Price, and Ivan Viola

IEEE Transactions on Visualization and Computer Graphics, Volume 18, Issue 12, pages 2265-2274, 2012
DOI: 10.1109/TVCG.2012.188

©2012 IEEE. Reprinted, with permission, from Veronika Šoltészová, Çağatay Turkay, Mark Price, and Ivan Viola, A Perceptual-Statistics Shading Model, IEEE Transactions on Visualization and Computer Graphics, December 2012. In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of TU Wien's products or services. Internal or personal use of this material is permitted.

A Perceptual-Statistics Shading Model

Veronika Šoltészová, *Student Member, IEEE*, Çağatay Turkay, *Student Member, IEEE*
Mark C. Price, and Ivan Viola, *Member, IEEE*

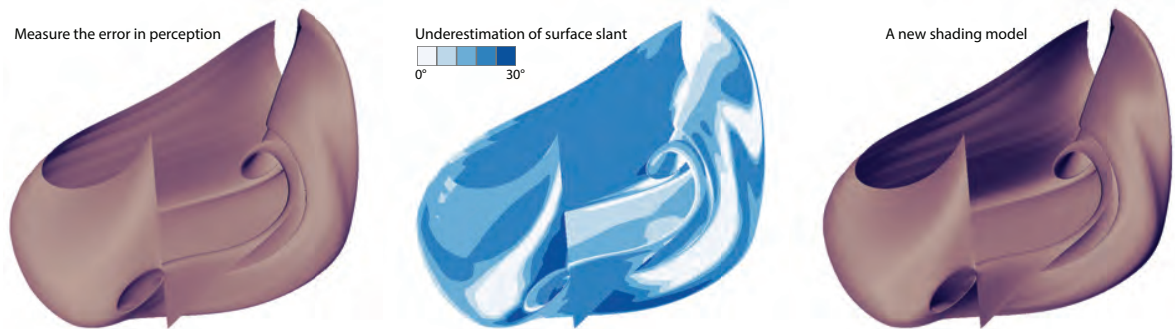


Fig. 1. The concept of iterative evaluation-analysis-redesign of a visualization technique is shown on a stream surface dataset. Analysis of the perceived surface slant while using a chosen shading model – the Lambertian shading model [9] on the left leads to a statistical model of the perceived error plotted in the middle. The statistical model of the error is then taken into account in the new shading model which aims to compensate for it, shown on the right.

Abstract— The process of surface perception is complex and based on several influencing factors, e.g., shading, silhouettes, occluding contours, and top down cognition. The accuracy of surface perception can be measured and the influencing factors can be modified in order to decrease the error in perception. This paper presents a novel concept of how a perceptual evaluation of a visualization technique can contribute to its redesign with the aim of improving the match between the distal and the proximal stimulus. During analysis of data from previous perceptual studies, we observed that the slant of 3D surfaces visualized on 2D screens is systematically underestimated. The visible trends in the error allowed us to create a statistical model of the perceived surface slant. Based on this statistical model we obtained from user experiments, we derived a new shading model that uses adjusted surface normals and aims to reduce the error in slant perception. The result is a shape-enhancement of visualization which is driven by an experimentally-founded statistical model. To assess the efficiency of the statistical shading model, we repeated the evaluation experiment and confirmed that the error in perception was decreased. Results of both user experiments are publicly-available datasets.

Index Terms—Shading, perception, evaluation, surface slant, statistical analysis.

1 INTRODUCTION

The major effort of computer graphics initially focused on the production of synthetic scenes that are indistinguishable from a photograph. From the visualization perspective, the *user-centric* aspect of rendering is more important than the *physics-centric*, and the focus is put on 3D scene understanding rather than on a physically-correct representation of a scene.

From the user-centric aspect, 3D shape and depth cues are important. Shape perception is mostly based on local features of surfaces, i.e., patterns of reflected light that are based on the surface orientation and the illumination direction, and texture deformation that is based on local curvature. Depth cues allow for correct depth ordering of structures and depth judgment. To resolve these cues, the visual system

uses not only stereopsis, perspective and kinetic cues but also our understanding of occlusion, shadows and haze. The judgment of depth is based on the global features of the scene while the judgment of shape considers mostly the local properties of the objects in the scene.

The user-centric aspect of rendering has been represented by styles that mimic techniques used in the craft of illustration. These techniques claim to be more efficient in terms of visual processing than a physics-centric representation of the same scene [12, 32]. Some rendering styles abstract from the realistic scene appearance by exaggerating the Lambertian shading gradient transitions [29]. Even though this approach has initially mimicked artwork, an increasing number of techniques are now motivated by new knowledge from vision research [37, 38]. Although perceptual evaluations of rendering techniques have been conducted in many recent reports, they have only rarely triggered a re-design of the original technique with the goal of perceptual improvement [13, 27].

The shading models mentioned above have an *imperative* character – an algorithm *dictates* the visual appearance that is displayed to the viewer. The viewer then extracts relevant information such as surface of objects, depth, and distances between them. The algorithm is independent of how accurately the intended information is conveyed. However, in contrast to the previous shading models, we present a shading model that starts as a classical *imperative* algorithm, but is then *declaratively* modified to improve the surface perception. This can be achieved through several iterations.

In this paper we first analyze the error of perceived surface orienta-

- Veronika Šoltészová and Çağatay Turkay are with the Department of Informatics, University of Bergen, Norway, E-mail: (Veronika.Solteszova | Cagatay.Turkay)@ii.uib.no.
- Mark C. Price is with the Psychology Faculty, University of Bergen, Norway, E-mail: Mark.Price@psysp.uib.no.
- Ivan Viola is with the Department of Informatics, University of Bergen and Christian Michelsen Research, Norway, E-mail: Ivan.Viola@ii.uib.no.

Manuscript received 31 March 2012; accepted 1 August 2012; posted online 14 October 2012; mailed on 5 October 2012.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

tions from shading, utilizing a common shading model (Figure 1 left). We perform statistical analysis on the data collected from a perceptual study that reveals systematic errors of human visual shape perception. This error, i.e., angular deviation between the ground-truth and perceived surface normals, is color coded and mapped to the stream surface in Figure 1 middle. From the statistical error description, we define a correction scheme. Next, we re-render the scene with a corrected rendering approach (Figure 1 right) and conduct another user study to analyze the new error trend. We propose a new concept of iterative modifications that allow the shading model to converge to a model with accurate perception where the distal and proximal stimulus match.

The major contributions of this paper are:

- a new concept: our work represents a next step in user-centric shading for scientific visualization that upgrades an imperative visualization algorithm with a declarative optimization, motivated by increasing the accuracy of perception,
- new knowledge: through perceptual evaluation we obtained new knowledge about error-distribution in shape perception according to the scene characteristics,
- a new shading model: we obtain a new shading model from the iterative evaluation and improvement concept that enhances surface shape perception,
- a publicly-available dataset which includes results of our experiment as well as the look-up map stored as a texture.

Previous approaches, even if they evaluated some perceptual error, did not use it for any improvement scheme, which is a part of our declarative concept. Our work presents a *missing link* in the visualization pipeline shown in Figure 2 in red which opens a new field of possibilities.

2 PREVIOUS WORK

For two millennia, scientists have been trying to elucidate the mechanisms in the human visual system (HVS) that are responsible for 3D shape perception. This topic remains an active area of multiple research disciplines such as psychology, neuroscience, computer science, mathematics, and physics. From the physics point of view, the sensory information is limited to patterns of light and is confined to their 2D projection on the retina. Using this sensory input, the HVS extracts information about the shape and the arrangement of objects with respect to their environment [34].

2.1 Perception of surfaces

The shape of an object is defined by the properties of its contour and its surface which does not change under similarity transformations. Despite the fact that the 2D retinal projection of the object depends also on its orientation relative to the observer, the percept of the shape tends to remain constant. This phenomenon is called *shape constancy* [28].

The HVS constructs a mental image of an object from a combination of top-down cognition and sensory input. At the lower sensory level, this includes the intensity variation of shading, texture gradients, edges and vertices. At the higher cognitive level, it includes salient features such as occlusion contours (object-background separation) [34]. Cole et al. showed that certain shape cues can be extracted solely from important lines, even though shape cues from shaded images are more accurate [6]. However, shading alone cannot yield the depth structure of a scene correctly [7]. The depth cues from shading are poor when compared to the retinal disparity (stereopsis) and kinetic cues [14].

Shading is specified by multiple parameters, i.e., the local surface reflectance properties, the angles between the surface normal and the direction of the light sources and the viewer. The judgment of shape is therefore a result of observers' assumptions regarding several parameters. The assumptions can vary between observers. Belhumeur et al. [2] introduced the term *bas-relief ambiguity*; when an unknown object with Lambertian reflectance is viewed orthographically, there

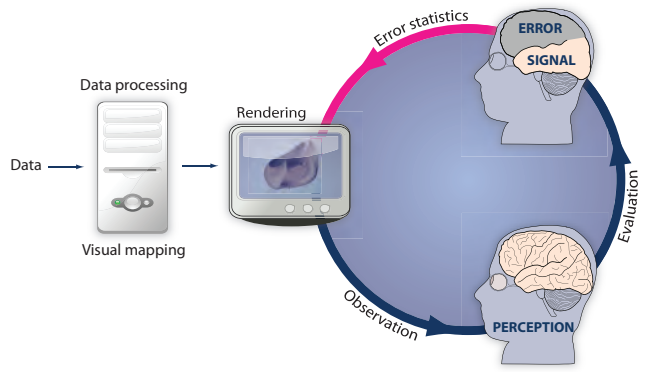


Fig. 2. The concept of iterative evaluation and design of a rendering technique. The original visualization pipeline contains no cycles and ends at the stage when the image is perceived by the user. The new concept contains a loop; The accuracy of perception is evaluated and the original rendering method is modified based on the measured error in perception.

is an implicit ambiguity in determining its 3D shape. For example, in a bumpy scene casting shadows, it is not possible to distinguish whether the light direction is more slanted or if the bumps in the scene are deeper. The object's visible surface $f(x,y)$ is indistinguishable from a generalized bas-relief transformation of the object $f(x,y) = \lambda f(x,y) + \mu x + \nu y$.

There is an evidence that the *pictorial relief*, i.e., the imaginary relief extracted from a 2D projection of a 3D scene, such as a rendering or a photograph, is systematically distorted relative to the actual structure of the observed scene [7, 34]. The variations among observers' judgments were revealed to be complex and thus could not be accounted for by a simple depth scaling transformation. However, subsequent analyses showed that almost all of the variance could be roughly accounted for by an affine shearing transformation in depth [34].

Mamassian and Kersten investigated the perception of local surface orientation on a simple smooth object, under various illumination conditions [21]. They analyzed perceived local orientations for several points on the surface and quantified the *slant* and *tilt* of the local tangent plane. By *slant*, we understand the angle between the surface normal and the view vector and, by *tilt*, the azimuth direction of the surface normal in the eye space [6]. This definition is illustrated in Figure 3. Mamassian and Kersten observed that slant was underestimated for slants larger than 20° and overestimated under this value. This systematic error in slant perception results from the lack of visual reference and indicates that relative slant is a more robust cue [11]. Because of the absence of binocular disparity [19] and environmental cues, such as the presence of a frame [36], the brain receives the information that the rendering is, in fact, flat. This information is in conflict with cues from shading and therefore, the mental image extracted from the rendering is flattened in a systematic fashion.

To resolve these ambiguities, the HVS tends to assume a certain light direction [24]. Johnston and Passmore suggested that the slant discrimination declined with rotation of the light direction vector towards the viewpoint [14]. Follow-up studies indicated that this direction is from above the viewer and 12° left from the vertical axis [33, 20]. O'Shea and colleagues studied the assumed slant of the light direction on purely diffuse surfaces with no shadows [26]. They demonstrated that the surface slants were most accurate when the light source was $20^\circ - 30^\circ$ above the viewer.

Mingolla and Todd [24] concluded that the HVS initially assumed Lambertian reflection on all surfaces. Furthermore, they suggested that the surface orientation was detected locally, and global shape was determined by smoothing over local features. Fleming et al. studied mirror-material surfaces, i.e., surfaces riddled with specular highlights that contained no shading [8]. They concluded that the HVS can somehow exploit specular reflections to recover three-dimensional

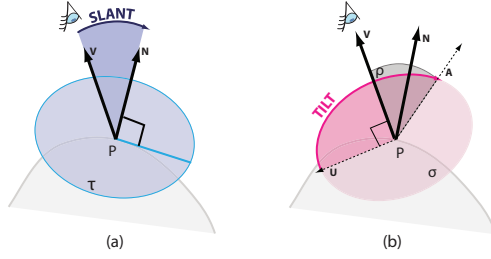


Fig. 3. The slant angle θ is defined as the angle between the surface normal \mathbf{N} at a point P and the viewing vector \mathbf{V} . τ denotes the tangent plane at P and \mathbf{U} the up vector of the viewer's coordinate-system. σ is a plane such that $P \in \sigma$ and $\mathbf{V} \perp \sigma$ and ρ denotes the plane defined by \mathbf{V} and \mathbf{N} . The tilt angle ϕ is then defined as the angle in the left-handed system between \mathbf{U} and $\mathbf{A} = \rho \cap \sigma$ in the halfplane (ρ, \mathbf{V}) defined by \mathbf{N} .

shape. The HVS treats specularities somewhat like textures, by using the systematic patterns of distortion across the image of a specular surface to recover 3D shape. Other studies also provide evidence about the influence of specular highlight on the perception of surfaces and demonstrate that the shininess of surfaces enhances the perception of curvature [25, 35].

Illustrators tended to exaggerate salient features such as curvature or important lines. Their methods have been mimicked by the graphics community. Exaggerated shading [29], geometry manipulation [15], light warping [37] and radiosity scaling [38] are good representatives. These techniques, however, were not derived from prior knowledge of a measured perceptual error. In contrast to prior work, we are presenting a novel concept where the visualization technique is based on a statistical model of the error in human perception. In particular, we target underestimation of surface slant of diffuse shaded surfaces. However, our concept can be applied to any self-chosen visualization technique that yields a measurable systematic error in perception.

2.2 Psychophysical experiments

The first experiments investigating human perception of 3D shapes were performed in the 19th century. The available information about these experiments is very poor, and therefore one should interpret their results with caution [34]. In the experiment of Mingolla and Todd [24], observers judged slants and tilts of numerous regions within shaded images of ellipsoid surfaces under varying illumination direction. The ellipses also had various shape, orientation and surface reflectance.

The works of Koenderink et al. [17] and Todd [34] describe the three most frequently employed experiments for probing perceived surfaces.

Relative depth probe task: Observers are exposed to a shaded surface. Two points on the surface are marked with dots of different colors. The observer is asked to choose which point he or she perceives closer in depth by pressing a dedicated key. Variations of this task were employed recently to assess visualization quality [18, 32].

Gauge-figure task: This task, designed by Koenderink et al. [16], allows one to determine the perceived orientation of a surface.

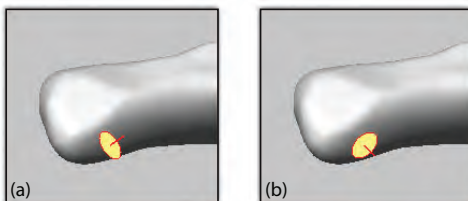


Fig. 4. Example of (a) a bad placement and (b) a good orientation of a gauge figure (red-yellow Tissot's indicatrix) over a shaded surface.

It uses a Tissot's indicatrix, i.e., an ellipse of distortion – a mathematical tool that characterizes distortions from a map projection. When the indicatrix is aligned with a surface that is perpendicular to the viewing direction, it appears as a circle. When the surface is slanted from the viewing direction, it is seen as an ellipse. A gauge-figure consists of a Tissot's indicatrix and a stick perpendicular to the plane defined by the indicatrix. On each trial, the observers' task is to align the indicatrix with the perceived shaded surface. At the same time, the stick should be aligned with the surface normal at the point where it intersects the surface. In Figure 4, we illustrate an example of a bad and a good orientation of a gauge figure. This task has been employed for example by O'Shea et al. to measure the accuracy of surface perception under varying slant of the illumination direction [26]. Šoltészová et al. utilized this test to compare the surface perception for different styles of shadow rendering [32].

Cole et al. conducted a large-scale gauge-figure experiment, where they compared the accuracy of surface perception from automatic and man-made line-drawing representations of objects compared to their fully-shaded renderings [6]. Their experiment is the most relevant for our work. Their study was performed on 14 different images, both organic and man made. On each object, they randomly selected 90, 180 or 210 positions. In all, they collected 275K solved gauge-figure trials accomplished by a total of 560 people and published this large dataset including user responses, datasets, scene settings and documentation.

Depth-profile adjustment mask: On each trial, observers are exposed to a shaded surface overlaid by aligned and equally spaced dots. In a second separate window, these dots are presented over a blank background and the observer is asked to adjust them so that they fit the perceived height profile defined by the dots in the first window.

Summary: Koenderink and colleagues compared these three tasks [17]. Coherent results can be achieved across observers and tasks. By far, the easiest and the most natural task to perform is the gauge-figure task. The judgment is instant, with no obvious reasoning; observers do not have to deduce their answers from their mental image. The pairwise depth-comparison task is also easy, but feels more boring and less natural. Observers have to abstract their answer from what they have perceived. It involves simple overt reasoning. The cross-section reproduction tasks feel not so much unnatural as indirect. With respect to reliability, the gauge-figure task is the most reliable.

3 PERCEPTUAL STATISTICS

In the original visualization pipeline, the data pass through the following stages until they reach the observer. After the acquisition stage, the data can be analyzed, filtered or processed in the data enhancement stage and later mapped to visual properties. Finally, the data are rendered and presented to the user. In some cases, the effect on perception is evaluated. Even though this is a step towards the perceptual aspect of visualization, the link from the evaluation back to the design of the rendering technique is practically non-existent.

In Figure 2, we show our new concept. We establish a new link that connects the results of an evaluation of a chosen rendering technique and the rendering technique itself. Starting from the rendering stage, the new pipeline now passes the following steps. The rendering is a distal stimulus which yields some sensory input which is interpreted by the HVS. This process is labelled perception. Evaluation refers to processing of the perceived information into the signal which corresponds to the ground truth and the error. Applying statistical methods to analyze the trends of the error allows us to model this error if it is systematic. This new knowledge is then sent to the rendering stage again. The rendering algorithm now becomes aware of the perceptual error it causes and can account for it.

If we see the pipeline shown in Figure 2 as a directed graph, the new link makes the graph cyclic. This allows for the possibility to loop between the rendering stage, evaluation and improvement. In this paper, we present how this concept can be used to improve the perception of surface slant in visualizations viewed on monoscopic screens, which is systematically underestimated [7, 34].

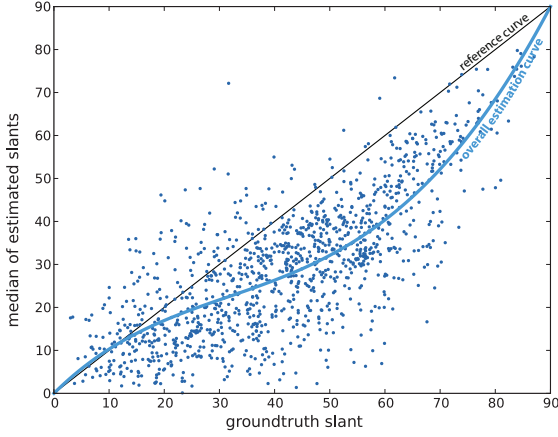


Fig. 5. Perceived surface slant as a function of the ground truth slant extracted from the dataset of Cole et al. [6]. Each dot represents the median of the entire set of trials at one sampling position. The *overall estimation curve* is a polynomial curve that is fitted to the data. The *reference curve* $x = y$ indicates a perfectly accurate estimation.

3.1 Analysis of the perceived surface slant

The perception literature reports that the surface slant, as deduced from monoscopic renderings of 3D objects viewed on a screen, is systematically distorted, however there is no model representing this phenomenon [26, 34]. The slant angle is understood as the angle between the surface normal and the viewing direction. We describe this effect with a mathematical model that was obtained through the statistical analysis of user responses. A model derived from statistical analysis of user evaluation has not been available before. It has been only attempted to model this effect as a parabolic function [26] or to use a simple shearing transformation in depth [34]. These approximations are consistent with the general expectation of perception but not founded on a statistical analysis of results of a perceptual study.

We obtained our model by analyzing users' responses collected as a publicly available dataset by Cole and co-workers as described in Section 2.1. The dataset contained results with fully-shaded and line drawing conditions. We analyzed only the responses for the fully-shaded condition. The line-drawing condition was completely excluded. For each of the 1200 sampling positions, we obtained the ground truth normal including the slant and the tilt angles and a corresponding set of normals estimated by the participants. In addition, for each sampling position, the authors of the dataset published the median of the corresponding set of estimates. They aimed to compare surface perception of 3D object representations on flat screens using monoscopic vision [6]. The overall dependency of the estimated surface slant θ_E and the ground truth θ_G slant is approximated with a polynomial fitting curve of the 4th degree and is shown in Figure 5. The *overall estimation curve* shows the trend of how humans tend to underestimate the surface slant. We originally computed different fitting curves with various specifications and obtained their goodness of fit (R^2 value) using the curve fitting tool of Matlab [22]. For various types of fit, we obtained the following R^2 values: Fourier fit of 1st degree – $R^2 = 0.773$, Fourier fit of 8th degree – $R^2 = 0.780$, exponential fit – $R^2 = 0.774$, cubic fit – $R^2 = 0.773$, and for polynomial fits of 4th degree – $R^2 = 0.775$, 5th degree – $R^2 = 0.775$ and 8th degree – $R^2 = 0.776$. As a trade-off between the complexity of the fit and the goodness of fit, we chose the polynomial fit of 4th degree.

However, the aggregated scatterplot in Figure 5 does not reveal a very interesting feature that is hidden in the dataset. We have separated the sampling positions into four groups according to the tilt ϕ of the ground truth normal: normals pointing upwards or north $\phi \in (315^\circ, 45^\circ]$; right or east $\phi \in (45^\circ, 135^\circ]$; downwards or south $\phi \in (135^\circ, 225^\circ]$; and left or west $\phi \in (225^\circ, 315^\circ]$. We define tilt (con-

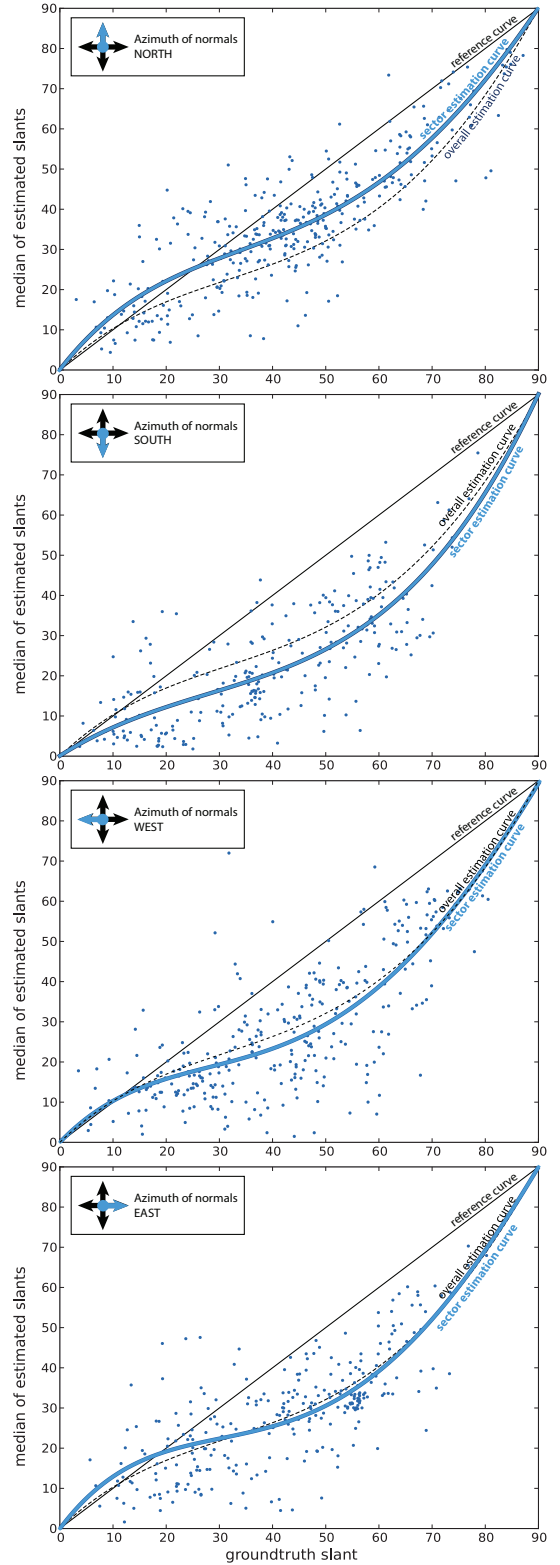


Fig. 6. Approximation of slant perception in four different sectors. The reference curve indicates a perfectly accurate perception while the sector estimation curves approximate the perception of slants in their respective sectors. We also plotted the overall estimation curve which indicates the average perception of slants in all sectors.

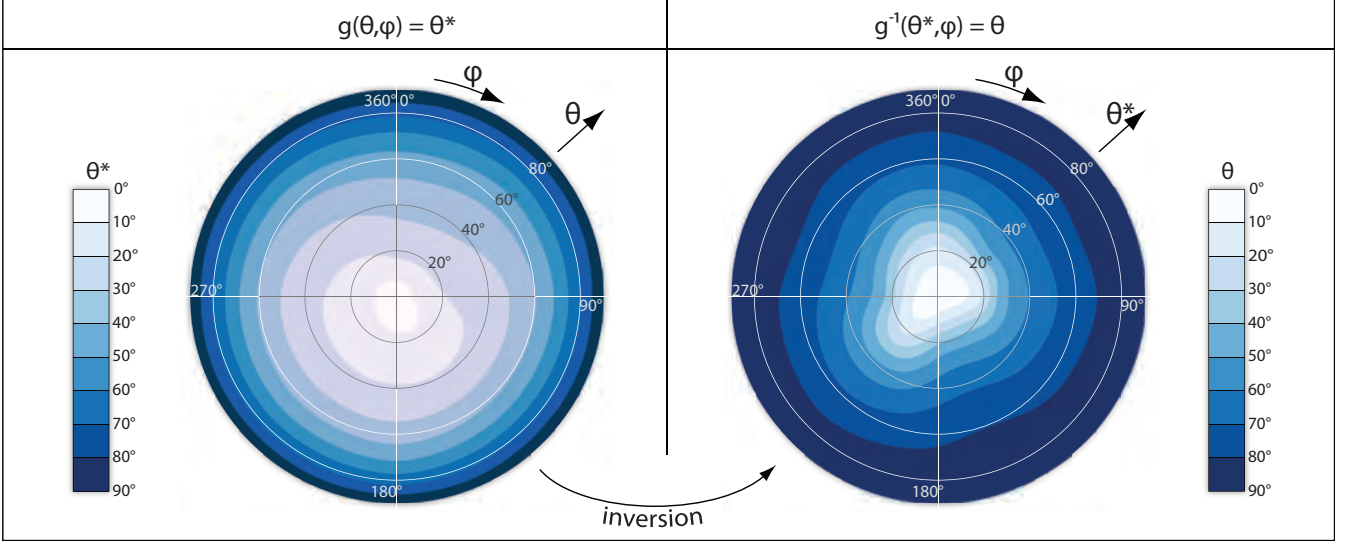


Fig. 7. Functions $g(\theta, \phi) = \theta^*$ and $g^{-1}(\theta^*, \phi) = \theta$ rendered as color-coded plots. Since $f = g^{-1}$, the right plot is also the look-up map which allows to efficiently find the slant angle θ of a normal which is perceived as θ^* .

sistently with the work of Cole et al.) as the azimuth angle on a compass where the wind directions are $N = 0^\circ$, $E = 90^\circ$, $S = 180^\circ$ and $W = 270^\circ$.

In Figure 6, we visualize the dependencies in each sector as scatterplots and fitted curves. The distribution and the *sector estimation curves* in the north and the south sector are very different. The slant of normals pointing north is underestimated less than average – the fitted curve is above the *overall estimation curve*. For the normals pointing south the situation is opposite. These slants are more underestimated than average – the fitted curve is under the *overall estimation curve*. The slant of normals pointing east and west are perceived very close to the average – the *overall estimation curve*. This finding is consistent with the statement of Todd that the underestimation of slant cannot be compensated by simple scaling in depth but by a shearing transformation in depth [34].

The crossing points of the *sector estimation curves* and the *reference curves* indicate the thresholds between over and underestimation of slant. In our results, these thresholds correspond to approximately $15^\circ - 25^\circ$ of the ground truth slant with the exception of the south sector. Mamassian and Kersten [21] expect this threshold to be approximately 20° which is consistent with our finding of $15^\circ - 25^\circ$.

We also considered a similar factorization of samples according to the maximal curvature (low, middle, high) but we did not find any remarkable dependencies between the error and curvature.

3.2 The model of surface perception

In order to model the human perception of slant, we compute a 2D map $f(\theta^*, \phi) = \theta$ which predicts that the slant angle of a surface normal should be θ so that it is perceived as θ^* . We divide the samples into bins that represent eight sectors: north, south, east, west, north-west, north-east, south-west, south-east. To obtain this map, we proceed as follows. For each sector, we calculate a polynomial fitting curve of the 4^{th} degree. Four of these sector curves (north, south, east, west) are plotted in Figure 6. These curves represent a function $g_\phi(\theta_G) = \theta_E$ which maps the ground truth slant θ_G in the sector ϕ to the estimated slant θ_E . For each curve, we set two boundary conditions: the curve must intersect points $(0,0)$ and $(90,90)$ since it is expected that the estimation of these boundary values is correct. These boundary conditions also guarantee that all curves start and end with the same functional value of θ_E and that the inverse function g_ϕ^{-1} is defined on the whole interval of slant $[0^\circ, 90^\circ]$. For g_ϕ^{-1} , the following condition holds: $g_\phi^{-1}(\theta_E) = \theta_G$. In other words, g_ϕ^{-1} predicts how the

slant angle of a surface normal should be so that it is perceived as θ_E and therefore $f(\theta^*, \phi) = g_\phi^{-1}(\theta^*)$.

So far, we have defined g_ϕ^{-1} for eight values of tilt ϕ only. In order to fill the missing values in the 2D map, we fit a smooth surface to the eight g_ϕ^{-1} aligned in polar coordinates according to their respective ϕ . To fit the surface, we used the surface fitting tool of Matlab [22]. Color-coded height maps of $g(\theta, \phi)$ and $f(\theta^*, \phi) = g^{-1}(\theta^*, \phi)$ are shown in Figure 7. The height map f , represented as a texture, allows for easy look-ups of the functional values of f at runtime. This texture is publicly-available for download [31]. While this texture is the best possible representation of our model, sometimes a functional approximation of $f(\theta^*, \phi)$ might be required. We found that \tilde{f} , which is a linear interpolation $g_{N=0^\circ}^{-1}$ and $g_{S=180^\circ}^{-1}$, yields very similar, however not identical, results. With g_N^{-1} and g_S^{-1} as polynomials of 4^{th} degree with coefficients $(5.77e-6, -1.19e-3, 7.3e-3, 0.11, 0.0)$ and $(4.21e-6, -6.73e-4, 1.88e-2, 1.69, 0.0)$ respectively, we define \tilde{f} as follows:

$$\tilde{f}(\theta, \phi) = \left| \frac{\phi - 180^\circ}{180^\circ} \right| g_N^{-1}(\theta) + \left(1 - \left| \frac{\phi - 180^\circ}{180^\circ} \right| \right) g_S^{-1}(\theta) \quad (1)$$

Ideally, the statistical model should be defined for each illumination algorithm individually because different algorithms might yield different response curves regarding the surface slant. We have obtained this model from renderings of objects from purely diffuse and opaque materials. The mathematical model could be different for specular and shiny or semi-transparent surfaces.

4 THE STATISTICAL SHADING MODEL

The shading information is one part of the sensory input which the human visual system uses for constructing its mental image of the 3D world. Indirectly, we are able to extract shape and deduce the surface normals from our mental image even though we are seeing only a 2D representation of an object, e.g., a photograph or a rendering on a computer screen. We have now analyzed and concluded that the surface normal we perceive is distorted from the ground truth normal of the depicted scene, and we have provided a mathematical model of this distortion. The difference between the ground truth and the estimated surface slant is mapped to a 3D model and plotted in Figure 1. Illumination algorithms used in computer graphics were until now unaware of this perceptual model. With this new input information, we propose a concept of how an illumination model relying on surface normals can be corrected so that the mental image is closer to the depicted scene.

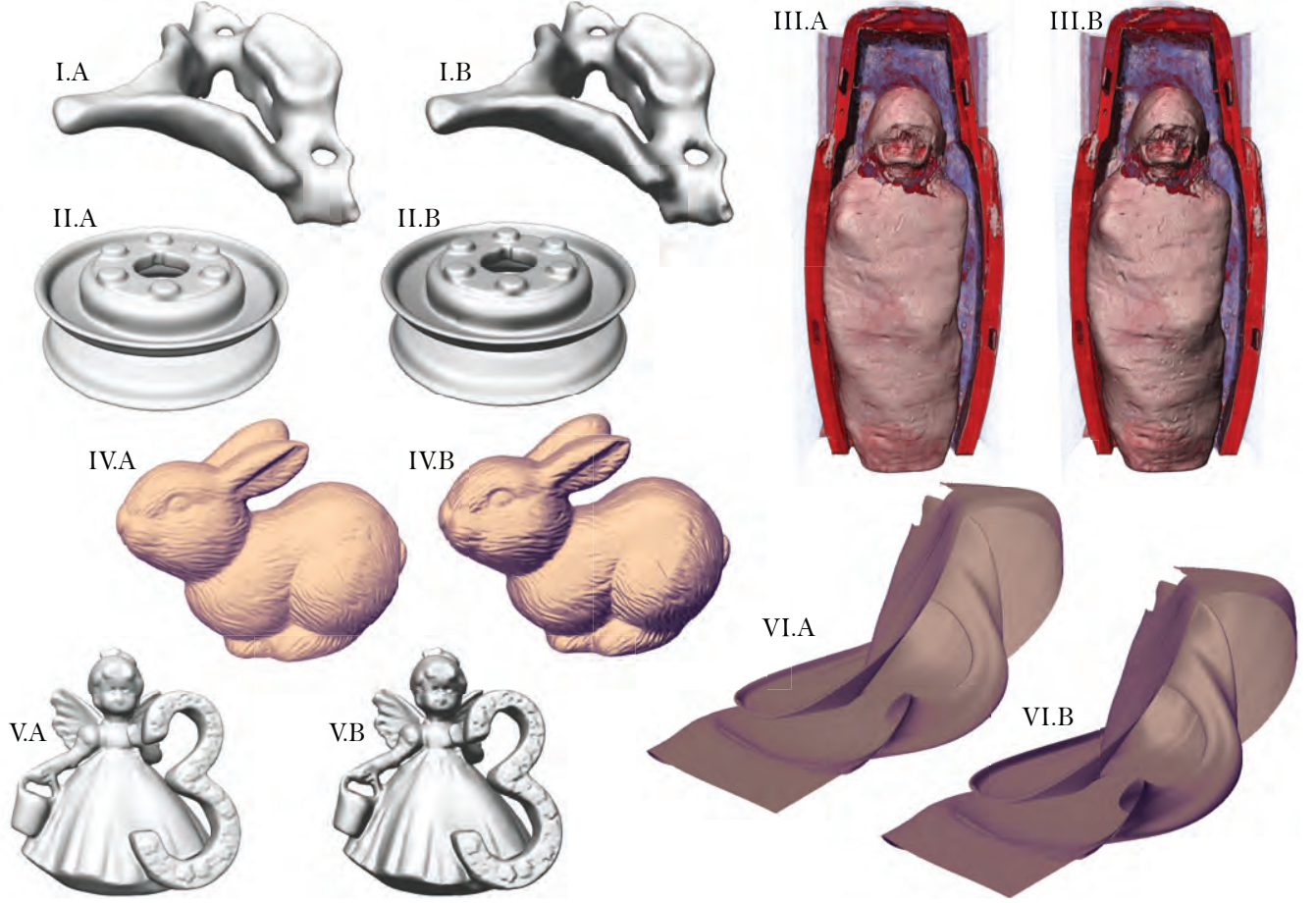


Fig. 8. The Lambertian shading using original normals (A) versus statistical shading model (B) shown on various datasets: I – cervical and II – pulley [6], III – a CT scan of a mummy, IV and V are geometry representations of laser scans of a bunny and an angel, and VI – a stream surface.

A rendering of a given scene geometry (distal stimulus) using normal-based shading, evokes its corresponding mental image (proximal stimulus) which can yield different perceived normals as those of the original geometry. Our goal is to match the distal and the proximal stimulus, i.e., to specify a shading model where the normals of the mental image and the ground truth normals match. We achieve this by manipulating the normals that are input into our shading model using a perceptual model corresponding to the original shading algorithm. In Section 3.2, we described how to obtain such a model and its approximating function $f(\theta^*, \phi) = \theta$. In our approach, we represented this function as a 2D look-up table stored as a texture where each pixel with coordinates (θ^*, ϕ) stores the value of $f(\theta^*, \phi) = \theta$. A color-coded representation of the look-up map and the coordinate system are shown in Figure 7.

A surface normal $n = (x, y, z)$ has slant θ and tilt ϕ given in projective space but is perceived to have wrong slant θ' . We shade the point with a modified normal $n' = (x', y', z')$ which has slant $\theta'' = f(\theta, \phi)$ and the same tilt ϕ . Notice that $\theta = g(\theta'', \phi)$, i.e., θ'' should be according to our theory perceived as θ . The components of the modified normal n' are then defined by the following equations:

$$x' = \frac{\sin(\theta'')}{\sqrt{x^2 + y^2}}x \quad y' = \frac{\sin(\theta'')}{\sqrt{x^2 + y^2}}y \quad z' = \cos(\theta'') \quad (2)$$

All illumination computation that follows is then executed with the new normalized surface normal $\frac{n'}{\|n'\|}$.

The concept of adjusting surface normals according to a given per-

ceptual model is applicable to any illumination computation scheme that is based on surface normals or gradients. To demonstrate the effect of our approach, we applied our model to Lambertian shading and used purely diffuse-reflective materials. In all settings, the light source conforms to the assumed light direction [26]. Figure 1 shows a stream surface before (left-most) and after our modification (right-most). Figure 8 contains more examples. A-images show the original shading with no modification of surface normals versus B-images showing our statistical shading. We included both datasets defined as volumes as well as geometry to show the general applicability of our technique. Objects I (cervical) and II (pulley) were also used by Cole et al. in their user experiment. Dataset III is a CT scan of a mummy visualized using gradient-based shading. Datasets IV and V were reconstructed from laser scans of a bunny and an angel. Dataset V is a geometry representation of a stream surface. All surfaces were shaded using Lambertian shading without (A) or with (B) modification of surface normals.

On first reflection, it might seem that similar results could be obtained by simply enhancing the contrast of the image [1] as in the juxtaposed images in Figure 9. Our method changes the intensities based on the surface normals, and therefore original intensities are mapped to a range of intensities. In contrast, global contrast enhancement maps the intensities to a single value. This shows that a global contrast-enhancement is a bijective function while the statistical shading is not. This effect is plotted in Figure 9. The standard deviation σ of global contrast enhancement is always zero which is not the case for statistical shading. Therefore, the same results cannot be obtained by simply enhancing the contrast of the image.

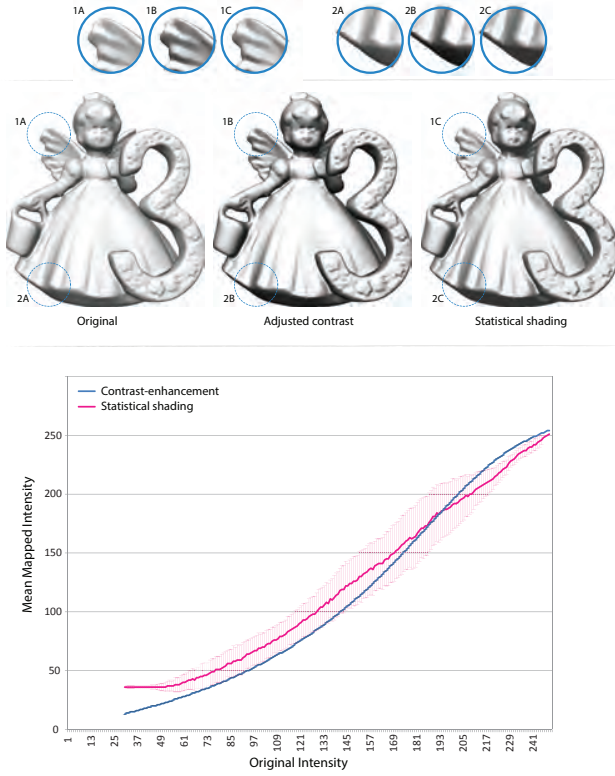


Fig. 9. Comparison of a contrast-enhanced image and a statistically shaded image. We plotted the mean mapped intensities of a contrast-enhanced image and statistical shading. The error bars represent the standard deviation σ of the mapped values.

5 VERIFICATION

Our hypothesis is that *the modification of normals causes the estimation of surface slant to be closer to the ground truth*. To obtain empirical support for our hypothesis, we studied perceptual judgments during the original shading condition (A) as opposed to our statistical shading condition (B). We then formally analyzed the difference in performance between the two conditions.

5.1 The Experiment

In order to measure the effectiveness of our technique, we conducted a new gauge-figure experiment. Instead of just relying on the results of the experiment of Cole and co-workers [6], we again tested condition A (original shading). This assured an appropriate control baseline, as we used a different rendering framework. Cole et al. generated their images with YafaRay which is a free raytracing engine [40] and defined their source of illumination as an environment map. We employed the commonly used Lambertian shading model and directional illumination.

We selected four distal stimuli from the experiment of Cole et al. – one organic dataset (cervical) and three man-made datasets (pulley, rockerarm, flange). Two of these stimuli are depicted in both shading conditions, in Figure 8 – I. (cervical) and II. (pulley). The stimuli were viewed on a flat computer screen using the same camera settings and viewport size as Cole et al. For each stimulus, we selected respectively 41, 42, 39 and 38 sampling positions for placing the gauge-figure from Cole’s dataset. The positions were heuristically selected from the whole set in the following way. For each object, the ground truth slants were best-possibly distributed over the interval $[0^\circ, 90^\circ]$ and the numbers of positions in each of four sectors (N,E,S,W) regarding the ground truth tilt were also balanced. In total, we used 160×2 distinct test cases: 160 gauge-figure placing positions and two shading conditions for each position. Each participant solved 2/3 of

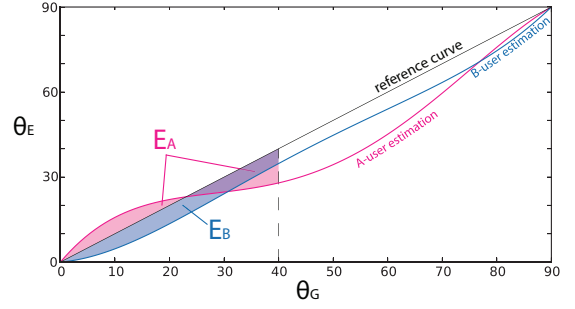


Fig. 10. The error areas of a selected participant for the original shading condition A – $E_A(0^\circ, 40^\circ)$ filled with pink and for our shading condition B – $E_B(0^\circ, 40^\circ)$ filled with blue.

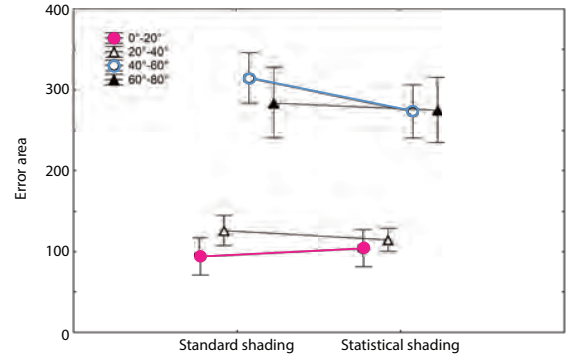


Fig. 11. Interaction plot between E for the two shading conditions: standard (A) and ours (B) in each of four subintervals of the curve. The vertical bars denote the 0.95 confidence interval. We found a significant improvement in the interval $[40^\circ, 60^\circ]$ – blue, a non-significant worsening in $[0^\circ, 20^\circ]$ – red, and non-significant improvements in $[20^\circ, 40^\circ]$ and in $[40^\circ, 60^\circ]$ – black.

all test cases so, in total, we collected at least 26 samples per test case and more than 8500 solved test cases overall. The collection of user responses is available for download [31].

Each of 40 participants attended two sessions. In each session he or she was tested on two pairs of stimuli with a 10 minute break between the pairs. The first pair of stimuli was presented in a different shading condition than the second. Half of the participants started with shading condition A and the other half with the shading condition B. The order was selected randomly in the first session, but in the second session, the order of shading conditions was reversed. For example, a random participant might be first presented with the stimuli cervical and pulley, and the shading condition A, then he had a short break to avoid fatigue and he continued with stimuli flange and rockerarm and the shading condition B. When this participant came to the second session, he started with the rockerarm, the flange, and shading condition A, and continued with the cervical, the pulley, and the shading condition B. The number of samples per position was balanced between participants.

We hired 40 participants for a financial compensation of 35USD equivalent for both sessions. The group of participants included 19 female and 21 male participants of 19 different nationalities. Participants were 21-47 years old but 87.5% belonged to the age group 20-30. Most of the participants were university students at the bachelor, master or PhD level. All of them had normal or corrected vision (lenses or glasses). 18 participants had skills with computer-assisted 3D tasks such as education in visual computing, mathematics or experience with 3D computer games. 37 participants worked on two different days. In three cases, the first session was in the morning and the second in the afternoon of the same day.

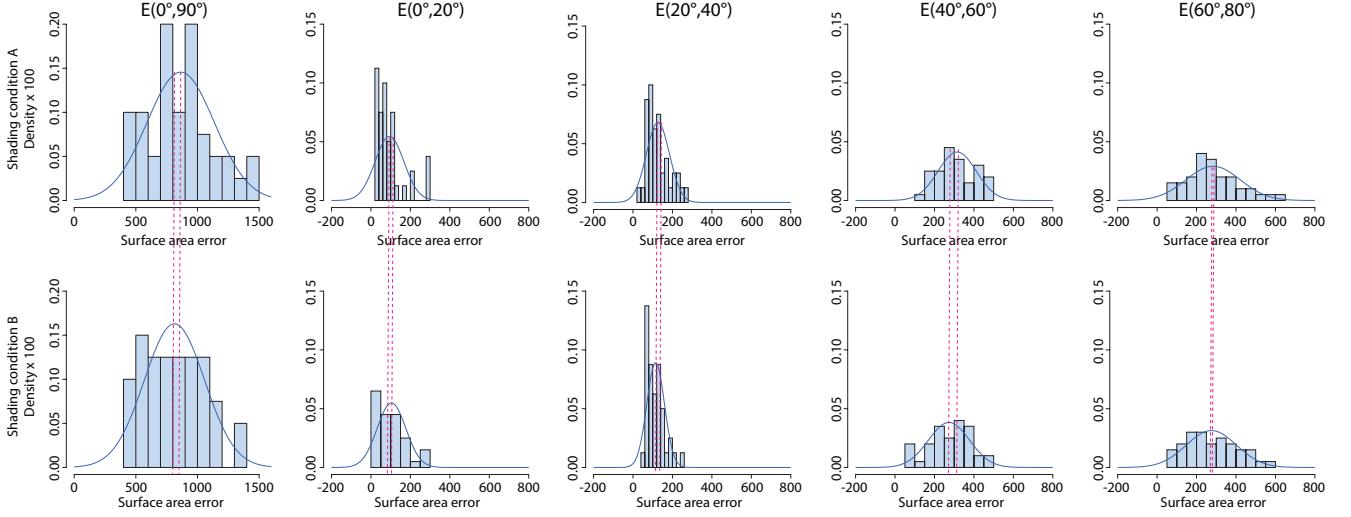


Fig. 12. Probability density of the surface area error E with histograms and approximative normal distribution curves for the entire interval $[0^\circ, 90^\circ]$ and four subintervals $[0^\circ, 20^\circ]$, $[20^\circ, 40^\circ]$, $[40^\circ, 60^\circ]$, and $[60^\circ, 80^\circ]$ and for both shading conditions. In intervals $[0^\circ, 20^\circ]$ and $[20^\circ, 40^\circ]$, the normality of the distribution is violated which can be deduced from the histogram. The orange dotted lines indicate the difference between the mean values of the shading conditions within the same interval.

5.2 Accuracy measurement of participants

To determine the accuracy of each participant, we approximated his or her responses for each shading condition (A and B) by two polynomial fitted curves of the 4th degree $f_A(\theta_G) = \theta_E$ and $f_B(\theta_G) = \theta_E$. θ_G and θ_E indicate the ground truth slant and the estimated slant respectively. Each curve was computed from at least 106 samples. We define the error measure $E(a, b)$ at an interval of slants $[a, b]$ as the area of the surface enclosed by the reference curve $R(\theta_G) = \theta_G$ and the user response curve $U(\theta_G) = \theta_E$:

$$E(a, b) = \int_a^b ||U(\theta_G) - R(\theta_G)|| d\theta_G \quad (3)$$

In Figure 10, we show the estimation curves of a selected participant for each shading condition – red for A and blue for B. The figure also illustrates the meaning of the surface area in a selected interval of slant angles (a, b) .

5.3 Analysis

To formally test whether the shading algorithm significantly improved participants' accuracy, we compared the error areas E between the two shading conditions A and B for each of the 4 intervals of the curve, i.e., $E(0^\circ, 20^\circ)$, $E(20^\circ, 40^\circ)$, $E(40^\circ, 60^\circ)$, and $E(60^\circ, 80^\circ)$. The division into subintervals was selected on a priori grounds. According to previous evidence [21] and also concluding from our own analysis, the underestimation of slant is zero at ca. 20° of ground-truth slant and highest for slants $40^\circ - 60^\circ$ (see also Figures 5 and 6). Hence we predicted different effects in each subinterval.

We conducted a 4×2 repeated measures ANOVA with the curve interval (4 levels) as one factor and the shading condition (2 levels) as the other factor. Due to violations of sphericity according to Mauchly's test, reported degrees of freedom and p-values are Greenhouse-Geisser corrected [10, 23]. The main effect of the curve interval was significant $[F(1.5, 59.8) = 68.4, p < 0.00001]$. A trend towards a main effect of the shading condition failed to reach significance $[F(1, 39) = 3.3, p = 0.08]$, although the area between ideal and obtained curves was numerically greater for the shading condition B (our new approach).

However, we obtained a significant interaction between the 2 factors, indicating that the beneficial effect of our shading algorithm differed for the different intervals of the curve $[F(1.8, 70.9) = 4.2, p = 0.02]$ as shown in Figures 11 and 12. Difference contrasts showed that a significant benefit $[F(1, 39) = 12.4, p = 0.001, r = 0.49]$ of the algorithm was obtained for the interval $[40^\circ, 60^\circ]$. Even though the error

bars indicating the 95%-confidence interval in Figure 11 do overlap, it does not imply that the effect is insignificant at 5% level [3]. For intervals $[0^\circ, 20^\circ]$, $[20^\circ, 40^\circ]$, and $[60^\circ, 80^\circ]$ respectively, $p = 0.35, 0.10$, and 0.56 . The effect of shading algorithm at the first 2 intervals was re-checked with non-parametric Wilcoxon tests [39] due to violations of normality for those distributions in a Shapiro-Wilk test [30], but still failed to show significant differences ($p = 0.23$ and 0.09 respectively). Figure 10 illustrates that the difference in surface areas between the two user estimation curves in the intervals $[0^\circ, 20^\circ]$, $[20^\circ, 40^\circ]$, and $[60^\circ, 80^\circ]$ is rather small compared to the interval $[40^\circ, 60^\circ]$ where the curves were expected to be further away from each other. Mean values and standard deviations of the error area distribution for each shading condition and for each interval of the curve are listed in Table 1.

In summary we found a highly significant effect of shading for angles in the interval $[40^\circ, 60^\circ]$. Moreover, in this curve interval, our shading manipulation had an effect size $r = 0.49$ that would normally be regarded as impressively large within the psychological testing literature [4, 5], accounting for 24% of data variance ($r^2 = 0.24$). Additionally, the significance level of this effect was high enough to exclude arguments that the effect was a Type I statistical error caused by multiple sampling at different intervals.

5.4 Discussion

Based on the results obtained in our gauge-figure experiment, we created and applied a second model of correction as described in Section 3.2. Rendering results of this iterative process of evaluation and re-design are illustrated in Figure 13.

We have shown that our modification of normals leads to more ac-

(a,b)	$E_A(a, b)$		$E_B(a, b)$	
	μ	σ	μ	σ
$(0^\circ, 90^\circ)$	863.82	274.06	813.34	244.78
$(0^\circ, 20^\circ)$	93.95	73.075	104.1862	72.59458
$(20^\circ, 40^\circ)$	126.34	58.52	114.82	44.6
$(40^\circ, 60^\circ)$	314.9	96.6	273.54	104.15
$(60^\circ, 80^\circ)$	284.08	136.8	275.38	127.06

Table 1. Table of mean values μ and standard deviations σ for the error area distribution within participants for each shading condition and each interval of the curve we analyzed.

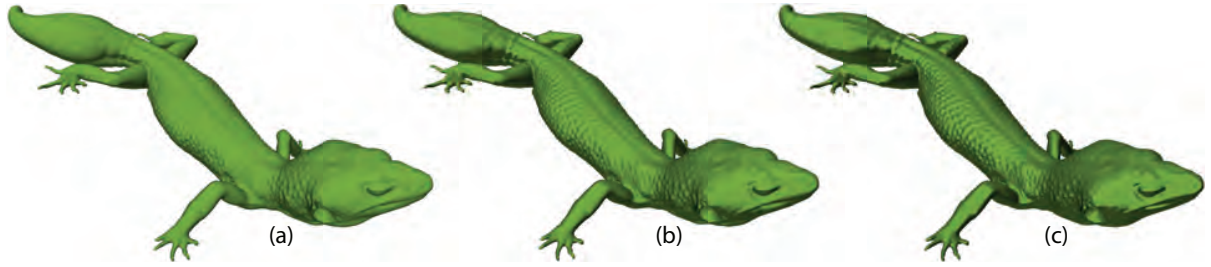


Fig. 13. Rendering results of a leopard gecko CT dataset of the iterative process *evaluation and re-design*: (a) the original Lambertian shading, (b) the result of a modification after the first user study, and (c) the result of a modification after the second user study.



Fig. 14. Comparison of (a) the Lambertian shading, (b) light warping, (c) exaggerated shading, (d) shearing along the z-axis and our approach using (e) the approximating function \hat{f} defined in Equation 1 and (f) the precise look-up texture to solve $f(\theta, \phi, \cdot)$.

curate perception of normals slanted $40^\circ - 60^\circ$. Our technique is not photorealistic. One could ask whether this is the case for other techniques that mimic methods from illustration and visual art? Were illustrators aiming to improve perception? We do not have access to a perceptual evaluation of other existing illustrative techniques such as light warping [37], and exaggerated shading [29]. In Figure 14, we juxtapose these to simple shearing along the z-axis, and with statistical shading in order to allow a subjective visual comparison. The two right-most visualization using the statistical shading model allow to compare the result of an approximative evaluation of $f(\theta, \phi)$ as defined in Equation 1 using function \hat{f} and precise evaluation using the lookup map.

6 CONCLUSION

We described a new concept of the visualization pipeline which allows one to update the rendering algorithm with new knowledge about how the human visual system misperceives the shape of 2D object depictions. Specifically, we studied the perception of surface slant of Lambertian-shaded surfaces and found a systematic distortion. We captured this effect as a function which predicts how the surface slant should be presented so that it is perceived as the ground truth slant. The function allowed us to modify the surface normals or gradients in the Lambertian shading model in a manner that was shown, via empirical testing, to objectively improve slant perception. Even though the trend for improvement did not reach significance when pooled over all slant values, we found a significant improvement in the interval $(40^\circ, 60^\circ)$ where the distortion of the slant perception is the highest.

6.1 Lessons learned

We found that the perception of normals pointing upwards in the eye space is clearly the most precise when compared to all other directions. Perception of normals pointing downwards is clearly the most inferior. Accuracy in the left and right directions is very similar. This characteristic of perception is illustrated in Figure 7 in the plot of $g(\theta, \phi)$. This shows that human ability to estimate surface slant is best on surfaces where normals point upwards and worst on surfaces where normals point downwards. We have not found a similar dependency of the estimation error from higher order surface derivatives such as curvature.

6.2 Limitations and future work

We studied the distortion of human surface perception using stimuli rendered with Lambertian shading of diffuse and opaque surfaces. Therefore, we cannot make a statement about this distortion if a different rendering algorithm, e.g., shadowing or ambient occlusion, was

to be used, or if the objects were to be made of a different material, e.g., semi-transparent or shiny. Each rendering algorithm and material should be studied individually and provided with a perceptual distortion model which is an inspiration for future research.

Since we have not evaluated the results after the second iteration, we are not able to say whether the iterations really converge to a perfect solution. Shape cues are not formed solely from shading. Even though shape extraction from a shaded image is more accurate, Cole et al. showed that certain shape cues can be extracted from line drawings as well [6]. Our method does not modify important lines since we are not deforming the objects. Therefore, we suggest that our method can be combined with a perception-enhancing geometry deformation in order to achieve the best results.

We would like to raise awareness that since we modify shading, we also modify luminance in the final appearance of objects. Since the depth perception is affected by the luminance channel [19], it might be worthwhile investigating how our modified shading influences the depth cues. Even though our aim was to show the effect on shape perception of (local) surface slant, we would like to encourage future studies of depth perception as well.

The manipulation of shading can influence the appearance of objects' material. The reason is that variations in shape tend to dominate variations due to shading [38]. This effect is visible in Figure 13. As we apply iterative modification of normals, the surface appears more shiny. This observation opens a new interesting direction of research to attempt to characterize a model that adjusts the cues from shading and contours while preserving the appearance of the material.

We observed that techniques that mimic illustrators' techniques are pursuing the same goal and, in our qualitative judgment, yield similar subjective effects. Speculatively, this suggests an intriguing hypothesis that illustrators used exaggeration of shading to better match the distal and the proximal stimulus.

ACKNOWLEDGMENTS

This work has been carried out within the IllustraSound research project (# 193170), which is funded by the VERDIKT program of the Norwegian Research Council with support of the MedViz network in Bergen, Norway (PK1760-5897- Project 11). The authors wish to thank the VisGroup at the University in Bergen, notably to Helwig Hauser, for discussions and to Herbert Grasberger who collaborated at scanning and creation of models IV and V displayed in Figure 8. Finally, the authors thank the anonymous reviewers for their constructive feedback and helpful comments.

REFERENCES

- [1] Adobe. Adobe Photoshop CS4 - The "Curves..." tool. <http://www.adobe.com/products/photoshopfamily.html>, 2008.
- [2] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.
- [3] S. Belia, F. Fidler, J. William, and G. Cumming. Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, 10(4):389–396, 2005.
- [4] J. Cohen. *Statistical power analysis for the behavioral sciences*. Routledge Academic Press, New York, 2nd edition, 1988.
- [5] J. Cohen. A power primer. *Psychological Bulletin*, 112(1):155–159, 1992.
- [6] F. Cole, K. Sanik, D. DeCarlo, A. Finkelstein, T. Funkhouser, S. Rusinkiewicz, and M. Singh. How well do line drawings depict shape? *ACM Transactions on Graphics*, 28(3):28:1–28:9, 2009.
- [7] E. De Haan, R. Erens, and A. Noest. Shape from shaded random surfaces. *Vision Research*, 35(21):2985–3001, 1995.
- [8] R. Fleming, A. Torralba, and E. Adelson. Specular reflections and the perception of shape. *Journal of Vision*, 4:798–820, 2004.
- [9] A. Gallardo. Lambertian shading. In *3D Lighting: History, Concepts and Techniques*, page 117. Charles River Media, Inc., Massachusetts, 2001.
- [10] S. Geisser and S. W. Greenshouse. On methods in the analysis of profile data. *Psychometrika*, 24:95–112, 1959.
- [11] J. J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.
- [12] B. Gooch, E. Reinhard, and A. Gooch. Human facial illustration: Creation and psychophysical evaluation. *ACM Transactions on Graphics*, 23(1):17–44, 2004.
- [13] D. H. House, A. S. Bair, and C. Ware. An approach to the perceptual optimization of complex visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):509–521, 2006.
- [14] A. Johnston and P. Passmore. Shape from shading. I: Surface curvature and orientation. *Perception*, 23:169–189, 1994.
- [15] Y. Kim and A. Varshney. Persuading visual attention through geometry. *IEEE Transactions on Visualization and Computer Graphics*, 14(4):772–782, 2008.
- [16] J. Koenderink, A. van Doorn, and A. Kappers. Surface perception in pictures. *Perception and Psychophysics*, 52(5):487–496, 1992.
- [17] J. Koenderink, A. van Doorn, and A. Kappers. Ambiguity and the mental eye in pictorial relief. *Perception*, 30:431–448, 2001.
- [18] F. Lindemann and T. Ropinski. About the influence of illumination models on image comprehension in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1922–1931, 2011.
- [19] M. Livingstone. *Vision and art – the biology of seeing*. Abrams, paperback edition, 2008.
- [20] P. Mamassian and R. Goutcher. Prior knowledge on the illumination position. *Cognition*, 81:B1–9, 2001.
- [21] P. Mamassian and D. Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351–2367, 1996.
- [22] MathWorks. Matlab: The language of technical computing. www.mathworks.com, 2012.
- [23] J. W. Mauchly. Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11(2):204–209, 1940.
- [24] E. Mingolla and J. Todd. Perception of solid shape from shading. *Biological Cybernetics*, 53:137–151, 1986.
- [25] J. F. Norman, J. Todd, H. Norman, A. M. Clayton, and T. R. McBride. Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision Research*, 46:1057–1069, 2006.
- [26] J. P. O'Shea, M. S. Banks, and M. Agrawala. The assumed light direction for perceiving shape from shading. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 135–142, 2008.
- [27] D. Pineo and C. Ware. Data visualization optimization via computational modeling of perception. *IEEE Transactions on Visualization and Computer Graphics*, 18(2):309–320, 2012.
- [28] Z. Pizlo and M. Salach-Golyska. 3-D shape perception. *Perception and Psychophysics*, 57(5):695–714, 1995.
- [29] S. Rusinkiewicz, M. Burns, and D. DeCarlo. Exaggerated shading for depicting shape and detail. *ACM Transactions on Graphics*, 25(3):1199–1205, 2006.
- [30] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4):591–611, 1965.
- [31] V. Soltészová. Perceptual-statistics web. www.iib.uib.no/vis/publications/publication/2012/Solteszova12APerceptual, 2012.
- [32] V. Soltészová, D. Patel, and I. Viola. Chromatic shadows for improved perception. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering*, pages 105–116, 2011.
- [33] J. Sun and P. Perona. Where is the sun? *Nature Neuroscience*, 1:183–184, 1998.
- [34] J. Todd. The visual perception of 3D shape. *Trends in Cognitive Science*, 8(3):115–121, 2004.
- [35] J. Todd and E. Mingolla. Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology*, 9(4):583–595, 1983.
- [36] A. J. van Doorn and J. Koenderink. The influence of environmental cues on pictorial relief. *Perception - ECVF Abstract Supplement*, 29, 2000.
- [37] R. Vergne, R. Pacanowski, P. Barla, X. Granier, and C. Schlick. Light warping for enhanced surface depiction. *ACM Transactions on Graphics*, 28(3):25:1–25:8, 2009.
- [38] R. Vergne, R. Pacanowski, P. Barla, X. Granier, and C. Schlick. Improving shape depiction under arbitrary rendering. *IEEE Transactions on Visualization and Computer Graphics*, 17(8):1071–1081, 2011.
- [39] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.
- [40] YafaRay. YafaRay 0.0.9: Yet another free raycaster. www.yafaray.org, 2008.

2.6 Perceptually Uniform Motion Space

Åsmund Birkeland, Çağatay Turkay, and Ivan Viola

IEEE Transactions on Visualization and Computer Graphics, Volume 20, Issue 11, pages 1542-1554, 2014
DOI: 10.1109/TVCG.2014.2322363

This is an open access article under the terms of the Creative Commons Attribution 3.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Perceptually Uniform Motion Space

Åsmund Birkeland, *Member, IEEE*, Cagatay Turkay, and Ivan Viola, *Member, IEEE Computer Society*

Abstract—Flow data is often visualized by animated particles inserted into a flow field. The velocity of a particle on the screen is typically linearly scaled by the velocities in the data. However, the perception of velocity magnitude in animated particles is not necessarily linear. We present a study on how different parameters affect relative motion perception. We have investigated the impact of four parameters. The parameters consist of speed multiplier, direction, contrast type and the global velocity scale. In addition, we investigated if multiple motion cues, and point distribution, affect the speed estimation. Several studies were executed to investigate the impact of each parameter. In the initial results, we noticed trends in *scale* and *multiplier*. Using the trends for the significant parameters, we designed a compensation model, which adjusts the particle speed to compensate for the effect of the parameters. We then performed a second study to investigate the performance of the compensation model. From the second study we detected a constant estimation error, which we adjusted for in the last study. In addition, we connect our work to established theories in psychophysics by comparing our model to a model based on Stevens' Power Law.

Index Terms—Motion visualization, motion perception, animation, evaluation, perceptual model

1 INTRODUCTION

THE use of motion can be seen in visualization techniques frequently. It has various purposes, for instance depicting a data attribute, to attract attention, or to convey shape information. An advantage of using motion in visualization is that motion detection is a pre-attentive process in the human cognitive system [1]. Motion can then effectively guide the users' attention to interesting features in the data and reveal small details in the motion pattern [2].

Flow visualizations are aimed to provide insight into how a fluid deforms under applied shear stress for a given situation. Unlike data representing solid physical objects, there are no *real* structures in a flow apart from the different flow patterns. Still, when referring to the topology of a flow, there exist abstract structures such as critical points and vortices. While topology is an important aspect of flow analysis, the velocity magnitude is also important in many cases.

Another reason for deploying animation of particles in visualization is that this is often a direct visualization of a particular phenomenon and might be therefore appreciated by the domain specialist. It exhibits qualitative characteristics such as detailed flow behavior, similar to the real observable phenomenon. One example can be seen in simulation of flow around an airplane wing. The lift from a wing is generated from the low pressure over the wing caused by the difference in velocity of the air moving above and below

the wing [3]. In medicine, blood-flow analysis is important in prevention, diagnosis, and follow-up monitoring of diseases. In this case, accurate flow data can be acquired with a range of techniques such as 4D Magnetic Resonance Imaging (MRI) for 3D flow, and B-flow ultrasound over time. For making correct decisions, multiple features of the flow have to be analyzed, such as pressure, vorticity and velocity.

Different visual cues can be utilized for different data attributes. Color coding is a typical method for depicting data. For instance, the velocity magnitudes in a flow field can be visualized with color. However, this provides no information regarding the direction of the flow, and other techniques such as glyphs must be applied in addition. The usage of color can be very efficient for a single data attribute. Still, there are other aspects of the flow that are often important for the user. An issue arises when more attributes should be visualized simultaneously. If color is already occupied for velocity magnitude, other means must be utilized.

In engineering, analyzing aerodynamics is often done by adding a flow of air around the object of interest. The air flow is as such not visible. To be able to actually see the flow, particles are added into the flow field. The motion of the particles conveys intuitively information regarding both the direction and the velocity magnitude of the flow. This technique has commonly been adopted into flow visualization [4], [5], [6], [7].

Investigations into how the human visual system processes motion can be approached from different directions. In the direction of neurophysiology, the fundamental laws of nature are applied for investigating the physical connection between motion detection and the neural activity. The approach is more directed towards signal processing, where one tries to understand how signals from the optic system are processed and transmitted into the visual cortex. This can be done either by applying physical models to simulate the neurological response in the brain, or procedures, where the actual signals in the brain are being monitored, for instance electroencephalography.

- Å. Birkeland is with the Department of Informatics, University of Bergen, Bergen, Hordaland, Norway. E-mail: aasmund.birkeland@gmail.com.
- C. Turkay is with the Department of Computer Science, City University London, Room: A304C, Northampton Square, London EC1V 0HB, United Kingdom. E-mail: Cagatay.Turkay.1@city.ac.uk.
- I. Viola's is with the Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11, E186, Vienna 1040, Austria. E-mail: viola@cg.tuwien.ac.at.

Manuscript received 27 Jan. 2013; revised 24 Mar. 2014; accepted 10 Apr. 2014. Date of publication 6 May 2014; date of current version 1 Nov. 2014. Recommended for acceptance by D.A. Bowman.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TVCG.2014.2322363

In psychophysics, investigation is performed by examining how physical stimuli can affect the perception in subjects. For motion detection, this is typically performed by providing a task to the user, based on a certain stimulus. Analysis is then based on examining correlations between the parameters in the stimulus and the response given by the user.

In this paper we have evaluated the perception of a speed-up factor of one motion pattern with respect to another motion pattern. We analyzed how the relative motion perception is affected by four distinct parameters: relative speed up factor between two sets of particles (speed multiplier), global scale of the velocities (the overall speed of particles moving across the screen), chromatic and luminance contrast, orientation, and direction of motion. In addition, we have tested for any influence by adding visual cues, in form of comet tails, and the point distribution (Poisson distribution versus random). We have performed user studies where we measured the subjects' estimation of relative speed between two separate sets of moving particles. We have discovered significant trends in estimation error for two of the parameters, speed multiplier and global scale of the velocities. The main contribution in this paper is the first compensation model for creating a perceptually uniform motion space, when using animated particles in visualization. From a series of perception studies we have shown how the compensation model successfully compensates for the effect of selected parameters. From our studies, we have shown how adding multiple visual cues, have a small improvement in perceived speed of animated particles. In addition, we provide statistical indication that the effect of chromatic versus luminance contrast are not as prominent on currently most widely used LCD monitors, as presented in previous work.

2 RELATED WORK

In order to understand how the human visual system is detecting and analyzing motion, there are several aspects that come into play. For instance, a typical view on problems in vision is that they can all be modelled as correspondence problems. Finding the correspondence of an object compared to the brain's representation of that same object is used in object recognition. In motion detection, the problem can be modelled as detecting correspondence over time. As a computational problem, motion detection can be seen as detecting changes in a given position. The *Reichardt detector* [8] is an implementation based on finding the correspondence over time.

In contrast to the Reichardt detector, an alternative model for motion detection involves finding change in luminance over time, known as the *temporal derivative*. A more detailed explanation can be found in related literature [9].

From a psychophysics perspective, we find much work in experimentation on the effect of different parameters when users evaluate motion. Experiments suggest that contrast change has an effect on the perception of the given speed [10], [11]. However, there is controversy regarding how this affects the perceived speed in general [12], [13].

Research indicates an impact of color and luminance to motion detection. A continuous change in luminance can

create apparent motion of stationary objects [14]. Derrington and Henning presented a study on the impact of color to the motion after-effect [15]. Backed up by the claim of color and motion being processed in parallel in the human visual system, evidence has indicated that applying only chromatic contrast compared to luminance contrast will cause a slow down in the perceived velocity [16], [17].

There exist rules which approach the subject of perception and how it scales to different stimulus. Weber's law is introducing the term *Just Noticeable Difference*, which means that the smallest difference between two stimuli is proportional to the absolute magnitude of the stimuli. Fechner's law states that the subjects' impression of a stimuli is proportional to the logarithm of intensity of the stimuli. In addition, there is Steven's power law [18], which is a more generalized description of the relationship between perception and stimulus. More details are found in existing literature [19], [20], [21], [22], [23].

Most work in motion detection results in qualitative statements that mainly explain the effect of a given stimulus. Little information is provided regarding how to compensate for any systematic distortions in the human visual system. In color theory, there exists a compensation model which creates a perceptually uniform color space called CIE 1976 (L^* , a^* , b^*) color space (CIELAB). In the context of motion, a similar model would aid visualization techniques based on moving elements in conveying the correct information according to the underlying data. A similar approach in the domain of 3D shape perception is taken by Šolteszova et al. [24]. In this work, the authors modify 3D shapes by building a statistical model of the error in perceiving shapes. Although in a different domain, this paper also demonstrates how visual representations can benefit from considering the perceptual aspects of the viewers.

We can explain our goal for this paper by drawing an analogy to color theory. As CIELAB is a color model where the perceptual difference for each step in the color space is uniform, we intend to create a compensation model for moving particles, which can compensate for systematic distortions in the visual system when estimating relative motion. This way, we can apply animated particles in visualization in a manner where the perceived information correlates better with the underlying data.

3 METHODOLOGY

In order to create a compensation model for a given set of parameters, we need to find the impact that each parameter has on the perceived relative motion. For estimating relative motion of animated particles, there are several parameters which can affect how a subject estimates motion. The density of particles, the size of the particles, contrast level between particle and background, contrast type (luminance or chromatic), relative difference in speed between particles, direction of motion, global scale of the velocities, shape of the particles, all can affect velocity estimation in different ways.

For our compensation model, we first selected a sub-set of the possible parameters to investigate. To assess a proper range for the parameters and the performance of our test-design, we performed a pilot study with a small set of

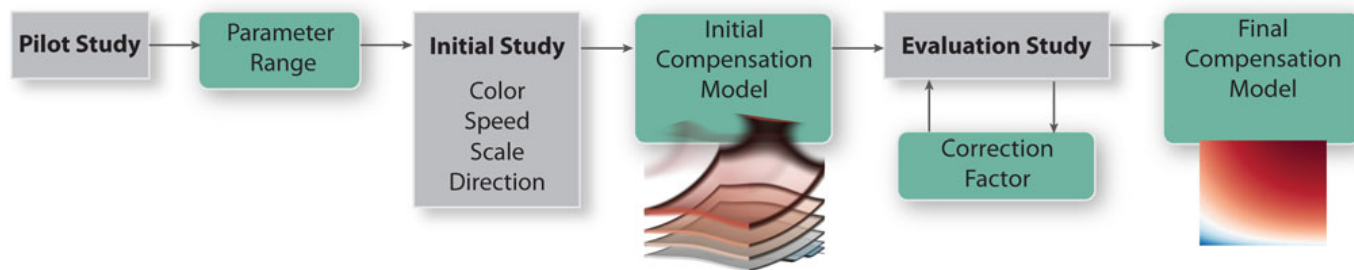


Fig. 1. Process pipeline for achieving perceptually uniform motion space. We started with a pilot study to determine suitable ranges in parameter space. From the pilot study, we designed an initial study to investigate the selected parameters. The output was an initial compensation model. In the evaluation stage, we performed two iterations to determine the efficiency of our model and adapted the model according to the new results.

participants. After adjusting the study based on observations from the pilot stage, we performed a larger study to find the trends for each selected parameter. Using the resulting trends we discovered in the initial study, we created an initial compensation model and started an iterative study process to test and refine the model. An illustration of the process is depicted in Fig. 1.

We started with a study using a simple setup, with basic particles with a uniform distribution and size. For a basic set of parameters, we chose four parameters that are seen frequently in visualization techniques using animated

particles. Examples of tasks involving each parameter can be seen in Fig. 2.

First, there is the range of screen-space velocities for the particles. When looking out the window of a fast moving car, it is difficult to clearly see stationary objects, like trees or road signs, close to the car. Based on this effect, we can deduce that there is a speed limit in the visual sector, for which humans can perceive objects clearly. We assumed that speed estimation would be affected by this effect as velocities would approach the limit of the visual system. To investigate this, we chose a parameter which would scale

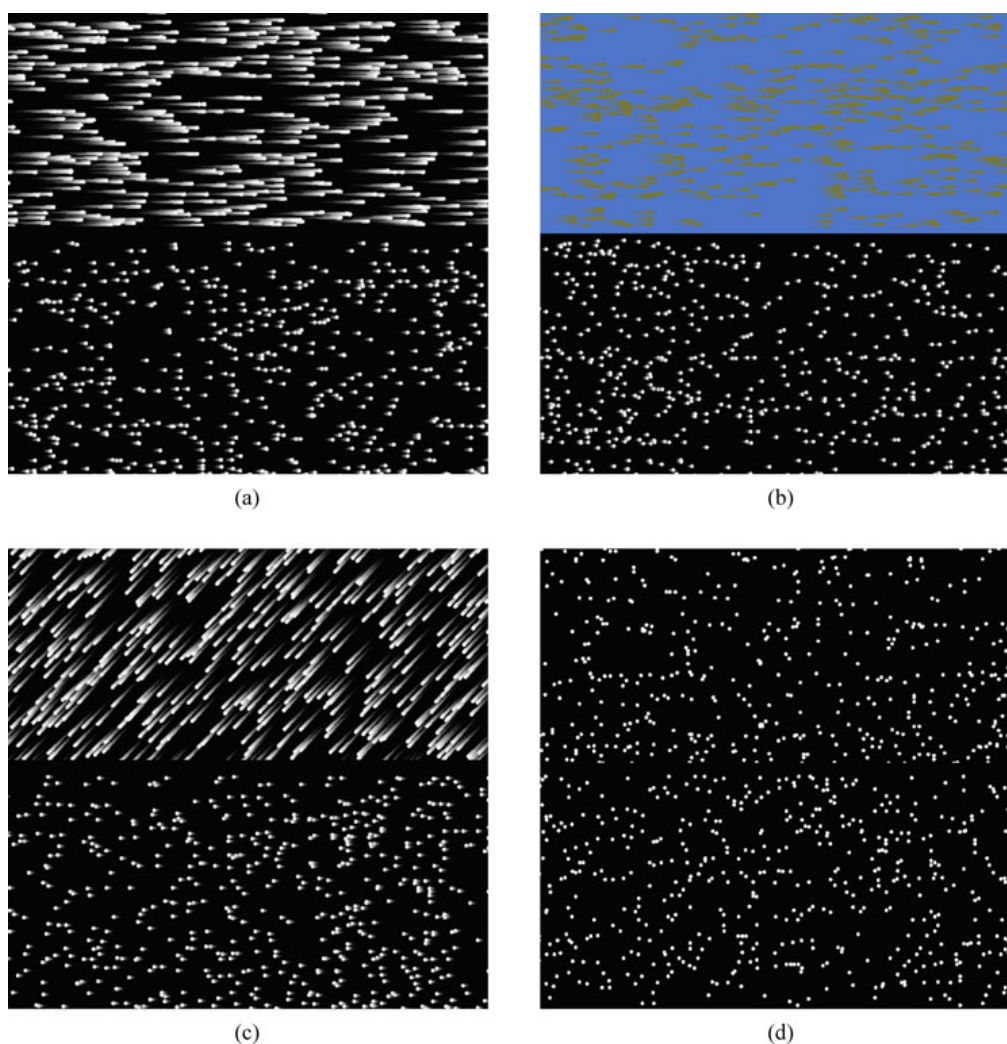


Fig. 2. Screenshots from tasks with different parameter settings. 2a shows a basic setup with direction angle set to 0 degree and the speed multiplier set to 5. 2b shows chroma only contrast. In 2c, the direction of the flow was set to 45 degree down and to the left. 2d shows the basic particles without a tail.

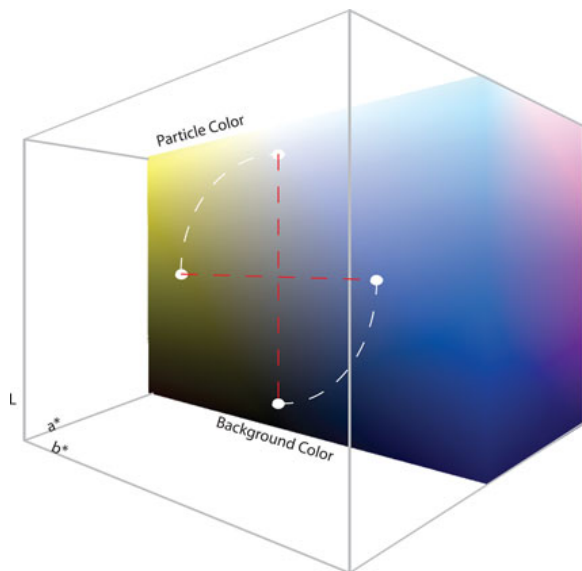


Fig. 3. The contrast-type parameter was set to be the angle of two points circling around a center point in CIELAB color space. 0 degree gave two colors with no chromatic contrast, and 90 degree returned two colors with only chromatic contrast, while always maintaining a constant contrast level.

up the speed of the animated particles globally, namely *global scale of velocities*.

Due to the asymmetries in horizontal and vertical vision [9] (e.g., due to the fact that our eyes are aligned horizontally), we draw the assumption that there would be a change in the perceived motion when viewing particles moving from left to right rather than moving from top to bottom. In addition, comparing motion when particles are moving in the same direction would be easier than comparing motion in the opposite direction. We chose to test for this effect by introducing a parameter which would set the direction of the particle motion. Direction ranged from 0 to 359 degree, where 0 degree meant the particles were moving in the same direction.

It has been stated that change in contrast and color can affect the perceived motion [17]. When mixing color coding and motion in visualization one should then be careful about the resulting contrast between the moving particles and the background. Based on the experiments presented in previous work, we assumed there would be a trend in estimation error as the contrast type would change from luminance-only to chroma-only. To adjust the contrast type, we generated colors by selecting points in CIELAB color space with luminance only. To shift the contrast type we rotated the points around the center-point within CIELAB color space up to 90 degree. 0 degree would then result in two colors with luminance-only contrast, and 90 degree would result in two colors with chroma-only contrast. The contrast level would remain the same for all degrees. Fig. 3 illustrates how the colors were selected.

As the difference in speed between the reference and test stimulus increases, we assumed that the perceived difference would not scale linearly. The main parameter tested was speed multiplier between two sets of moving particles. This corresponds to the theory of *Just Noticeable Difference*, meaning that the smallest difference we can detect, is connected to the intensity of the certain input. If the absolute

speed of the particles is high, the absolute difference between them should be comparably high.

We investigated the effect of adding additional visual cues, which would indicate the speed of the particles. To test if additional visual cues improved the perceived speed, we added comet tails to particles for half of the tasks generated. The length of the comet tails were linearly scaled by the speed of the particles and had a linear drop-off in width and opacity. We also tested for any influence of the distribution of the points drawn on the screen. A short pilot study was performed where half of the tests used a Poisson distribution and the other half used random sampling.

While contrast type and direction has a natural limit of range, the range of global scale and speed multiplier has no such limit. However, it is natural to assume that there is a certain limit for these parameters, where the error in estimation becomes too large to be clearly connected to the underlying data. In the pilot study, we found that estimating speed for particles moving with a velocity over 41.5 degree/second in the visual sector, the deviation in error becomes so large that any estimation from a user becomes meaningless when linked to the data. In addition, if the velocity magnitude of one set of particles became more than six times the velocity magnitude of another, the error in estimated speed multiplier became too high to be useful. We set the base speed to 0.83 degree/second, which constitutes one pixel per frame at 30 frames per second. This prevents any jitter in the movement of the particle. With a base level of 0.83 degree/second, the range of the global scale parameter was set to 1.0 to 7.0. With a maximum speed of 41.5 degree/second, we set the range of the speed multiplier to range between 1.0 to 6.0.

4 TEST DESIGN

The aim for the study was to test for perception of relative speed between particle sets. In particle based visualizations, there is a multitude of variations in complexity among particle types, size, density, direction of motion, flow topology and more. To investigate the perception of relative motion, we made a deliberate decision to start with a basic setup, which would test only one parameter at a time. This way, we can remove any unforeseen effects from other parameters.

In our tests we have two sets of particles displayed on the screen. One set is displayed at the top half of the screen (*test particles*) moving at a velocity \mathbf{v}_t , and one set is displayed at the bottom half (*reference particles*) moving at a velocity \mathbf{v}_r . The user was then asked to provide the multiplication factor, s , which satisfies the equation:

$$|\mathbf{v}_r| \cdot s = |\mathbf{v}_t|. \quad (1)$$

We have chosen to design the stimulus as a juxtaposition of two stimuli [25]. An alternative would be a superposition of the two stimuli, but this design will suffer from particle mixture and mutual occlusion. In such a case additional cognitive load will be required to isolate these two stimuli from each other, prior to the comparison.

The juxtaposition design has one problem though. When the two stimuli are far away from each other, the subject has

to frequently move the eyes to switch the focus. Therefore, in our study we have placed the stimuli as close to each other as possible. This allows the foveal and parafoveal vision to take part on the visual processing, without the need of frequent refocusing from one stimulus to another one.

Change in density as well as a repeating pattern can distort the perception of speed. To compensate for this, the particles for each half were generated randomly under the constraint that the average spatial distribution remained uniform regardless of velocity and direction similar to the experiments performed by Geesaman and Qian [26] or Watamaniuk et al. [27].

However, the random generation of particles might lead to structures in the motion pattern, i.e., several particles clustered together due to randomness in the distribution. To investigate whether such structures affect the test results, we perform an initial pilot study that compares our randomly distributed particles with those that are regularly distributed. In order to create moving particles that are regularly distributed, we generate our particles following a Poisson Disk distribution as suggested by McCool and Fiume [28]. We performed a small test with three subjects, where each subject completed 100 tasks. We then check whether there is any significant difference when particles are regularly distributed, we observed that the estimation error with the regularly distributed particles was slightly higher, i.e., average error -0.0416 with random versus -0.0991 with regular points. However, this difference was found to be insignificant with a p -value of 0.3412 when a regular two sample t -test was applied. Due to this result and due to the slightly better performance in error, we continue our study with randomly generated particles.

We selected subjects from various ages, gender, and professions. Each user was given a set of 100 tasks. To avoid any learning curve, we excluded previous participants for the consecutive experiments. The subjects were not informed of the parameter intervals and were told to estimate velocities based on visual impression and not from explicitly timing particles' traversal over the screen (for example by counting seconds and comparing distances). For the four rounds of user studies, we invited 10 participants for the pilot study, 22 for the initial round, 10 for each of the evaluation rounds.

For every test subject we used a 24-inch screen with a 16:10 aspect ratio. The pixel size was 0.27 mm and the canvas dimension was set to 768×768 . The user's head was approximately 50 cm away from the screen. While traditional CRT screens are typically used for experiments, we used modern LCD screens, which have sufficient quality for our experiments [29] and, more importantly, are utilized nowadays, in contrast to the CRT monitors that are practically not in daily use any longer. Wang and Nikolic stated that the use of LCD can in some cases be preferred, but not when the image changes rapidly. For our case, the smooth motion makes it beneficial to use LCD. In addition, the changes are not rapid. A small blur will appear from the points, but similar to the comet tails, this has a very little effect to the perceived stimulus as are shown in the results in Table 3. Another challenge with LCD screens is representation of color. While the LCD screens might not provide an absolute iso-luminant contrast, our motivation is driven by

real world application, and in a real world scenario, optimal conditions are never fully achieved. In most real world situations, the user will use an LCD screen with default settings.

If the eye is fixated on a continuous motion over some time, the eye will *adapt* to the motion. The reason for this is an effect called neural adaptation [30], where the neurons coding the particular motion reduce their responses. This can result in a distortion of our test. To avoid direction fatigue for the base speed, we alternate the direction of the flow for every other test. Furthermore, the subjects were asked to take a two-minute break after each 25 tasks have been completed.

The test design remained unchanged for the three iterations of testing.

5 USER STUDIES

5.1 Creating the Compensation Model

In the first round we wanted to investigate the effect of the selected parameters separately. We generated three types of tasks, one which tested the effect of global scaling of the velocities. The scaling is added on both sets of the particles. This changes Equation (1) to:

$$|\mathbf{v}_r| \cdot s \cdot \sigma = |\mathbf{v}_t| \cdot \sigma. \quad (2)$$

The range of σ was 1.0 to 7.0 and the range of s was 1.0 to 6.0 . For all the tasks involving global scale, the color-parameter remained constant at 0 degree and the motion direction remained strictly horizontal.

The second type was aimed for contrast type. We generated tests where colors were selected with the scheme explained in Fig. 3. The parameter ranges of 0 to 90 degree in CIELAB color space. The range of s was 1.0 to 6.0 . The global scale parameter σ was constant at 1.0 and direction remained strictly horizontal.

Finally we tested for the direction of the flow in the range of 0 to 359 degree direction. For each task we queried the subjects for speed multiplier between the two sets of particles in the range from 1.0 to 6.0 . The global scale parameter σ was set to 1.0 and color-parameter was set to 0 degree.

For each parameter tested, we created tests with random configurations within the given parameter space. For instance for contrast-type, a configuration might be as follows: $\sigma = 1$, $s = 2.3$, color-parameter = 90 degree and direction = 0 degree. To prevent samples from being too clustered in the parameter space, we constrained a random function to keep the same number of tests within each interval. In total 2,220 tasks were generated for the initial study. When the study was completed, outliers were removed. We utilized a visual inspection of the results supported by the Mahalanobis scores computed for each sample. Here we take a purely data-driven outlier-removal strategy, rather than considering specific participating individuals or the contextual properties of the test. In this stage, we compute a Mahalanobis score for each sample using the speed multiplier versus signed estimation error. The threshold for outlier removal was 14.2 , determined by visual inspection of the scores. This resulted in the removal of 22 outliers. The final sample size was then 2,198.

TABLE 1
Results of the General ANOVA Test Treating Each
of the 10 Bins as a Separate Group

Parameter	F-value	p-value
Global Scale	9.9745	1.440×10^{-13}
Direction	0.7131	0.490
Speed Multiplier	8.6012	9.176×10^{-13}
Chroma vs. Luminance	0.8376	0.5816

The high p-value show that there are no significant trend in estimation error compared to direction and contrast type.

From the results of the initial study we examined the relation between signed error and the parameters, i.e., if test-subjects overestimate or underestimate the speed. The trend in signed error is more relevant than unsigned. As there is a general trend in overestimation, we can compensate for this by slowing down the test particles. By this we will achieve a closer match between human perception of relative speed and the intended information communicated via the visualization.

We investigated the estimation trend for each parameter. The significance of the trends was evaluated by binning samples in parameter intervals. To provide a more robust significance evaluation, we performed statistical tests for different parameter intervals. In the following, we discuss the tests with 10 bins for each parameter. We start by applying a one-way Analysis of Variance (ANOVA) test by treating each bin as a separate group. When we observe the results for this general multiple group test, we see that there is a significant difference between the groups for the global scale and the speed multiplier parameter. However, for the direction and chroma parameters, we observe no overall significant difference between the multiple groups (refer to Table 1 for the corresponding results). In order to achieve more detailed results to explain the trends within the parameter intervals, we perform post-hoc tests. We prefer to do a *two-sample unpaired t-test with unequal variances* between each bin and the bin at the initial parameter setting. We follow this strategy (i.e., initial bin versus the others) since the parameter intervals are ordered and we want to investigate the trends in relation to this ordering. Moreover, we observe that the variance of the data is not equal within different intervals so we assumed unequal variances in our tests. Table 2 displays the t-scores for each parameter. The table shows the results from having 10 bins for each parameter. Interval shows the bin size in parameter space. N is the number of samples in the bin. Mean shows the average in signed estimation error and T-score shows how the bin compares to the top row. DoF indicates the degrees of freedom for the t-test. Since the different intervals contain samples from the same individuals under different conditions, i.e., varying parameters, we calculate the degrees of freedom accordingly and use $n/2 - 1$ as the formula where n is the total number of observations in the both groups. One point to mention here is that we checked for normality (using a Shapiro-Wilk Normality Test) on each of the bins prior to performing the tests whether to use non-parametric tests instead. We observed that for some of the bins, the

TABLE 2
Results from the *Two-Sample Unpaired*
T-Test with Unequal Variances

Direction				
Interval	N	Mean	T-score	DoF
0 - 36	150	-0.126	0.000	149
36 - 72	141	-0.155	0.682	145
72 - 108	142	-0.110	0.352	146
108 - 144	148	-0.094	0.722	148
144 - 180	145	-0.158	0.729	148
180 - 216	149	-0.220	2.054	148
216 - 252	146	-0.170	1.058	147
252 - 288	145	-0.107	0.424	148
288 - 324	128	-0.106	0.440	138
324 - 360	144	-0.155	0.745	146
Global Velocity Scale				
Interval	N	Mean	T-score	DoF
1.0 - 1.6	40	-0.206	0.000	39
1.6 - 2.2	34	-0.007	3.903	36
2.2 - 2.8	35	0.244	5.834	36
2.8 - 3.4	31	0.304	6.435	34
3.4 - 4.0	37	0.375	6.023	37
4.0 - 4.6	33	0.313	6.369	35
4.6 - 5.2	32	0.425	5.468	35
5.2 - 5.8	36	0.594	7.804	37
5.8 - 6.4	28	0.530	8.643	33
6.4 - 7.0	35	0.617	8.644	36
Speed Multiplier				
Interval	N	Mean	T-score	DoF
1.0 - 1.5	224	0.032	0.000	223
1.5 - 2.0	213	0.009	0.676	217
2.0 - 2.5	217	0.011	0.571	219
2.5 - 3.0	221	-0.016	1.279	221
3.0 - 3.5	216	-0.015	1.172	219
3.5 - 4.0	228	-0.099	3.464	225
4.0 - 4.5	187	-0.135	5.158	204
4.5 - 5.0	263	-0.112	4.165	242
5.0 - 5.5	207	-0.153	5.400	214
5.5 - 6.0	222	-0.200	7.144	222
Chroma vs. Luminance				
Interval	N	Mean	T-score	DoF
0 - 9	35	-0.104	0.000	34
9 - 18	47	-0.127	0.344	40
18 - 27	48	-0.133	0.405	40
27 - 36	40	-0.154	0.772	36
36 - 45	39	-0.176	1.164	36
45 - 54	39	-0.052	0.712	36
54 - 63	42	-0.102	0.032	37
63 - 72	38	-0.059	0.572	35
72 - 81	50	-0.145	0.681	41
81 - 90	41	-0.156	0.739	38

A significant correlation can be seen for the speed multiplier and global velocity scale parameters. For the direction parameter only a weak correlation can be seen, at around 180 degree. Unlike previous work, we could not detect any significant trend from change in contrast type. Here the results are collected into 10 bins for each parameter. Interval column indicates the parameter range for each bin. N is the number of samples in each bin. Mean shows the average in signed estimation error and T-score shows how the bin compares to the top row. DoF indicates the degrees of freedom for the t-test, where it is calculated as $n/2 - 1$ (n being the total number of samples) to account for repeated measurements.

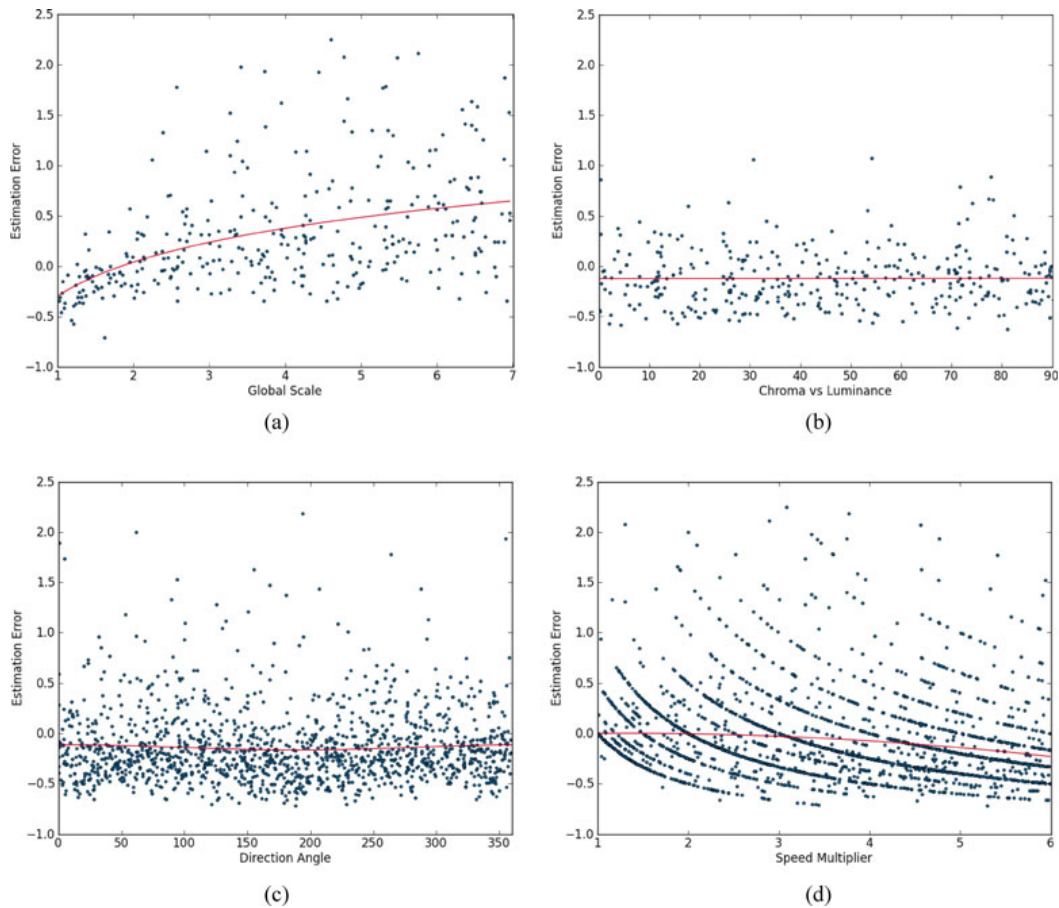


Fig. 4. Scatter plots of the samples generated from the first perceptual study round. The red curve shows the regression curves used for the compensation model. The curve for chroma-type (4b) shows no specific trend and thus this parameter was not included in the compensation model.

normality condition was not met. However, we still prefer the parametric t-tests since they provide us an insight on how the error evolves over the consecutive intervals.

We assumed that there would be a trend in estimation error which correlated with the relative motion between the two sets of particles. From Table 2, we can see there is a trend to continuously underestimate the speed value as the speed multiplier is increasing. Fig. 4d shows a scatter plot of the error in estimation compared to the speed multiplier parameter.

The results from the initial study also justify the assumption that there would be a change in estimation as the global scale of the velocities increases. Unlike the speed multiplier, subjects move from an underestimation to an overestimation. However, the slope for the regression curve for the relative motion factor remained unaffected by the global scale parameter.

We had a hypothesis that comparing particles moving in different directions would be more difficult than comparing particles moving in parallel to each other. The difference in direction, however, seems to have a very small impact on the estimation error. From the results we could only find a significant correlation in unsigned error. There was, however, a weak correlation in the signed estimation error. This leads to the conclusion that direction should be accounted for in our compensation model.

Based on previous experiments in prior work [17], we expected to see a general trend in underestimation when using iso-luminant contrast. However, we were unable to detect any *slow-down* effect in our experiment. Fig. 4b

shows the impact on the estimation error related to the contrast-type.

In addition, Table 3 shows the measured effect of the added visual cue. The effect is small and only prominent for animated particles with higher velocities.

The overall goal was to detect which parameters caused a trend in estimation error, which in turn, could be compensated for in particle-based visualization. From analyzing the result, we found strong correlations between speed multiplier and scale when comparing them with the estimation error. We also found a weak correlation between direction and estimation error. Therefore we chose to include the three parameters into our compensation model. We also checked for dependence between these parameters using

TABLE 3
This Table Shows the Mean Error and Standard Deviation of the Estimation Error for Each Task Type, Compared to the Overall

Parameter	Tail	Mean Error/Std. Dev.
All	without	0.335 / 0.283
	with	0.299 / 0.271
Global scale	without	0.489 / 0.489
	with	0.375 / 0.432
Contrast type	without	0.253 / 0.182
	with	0.250 / 0.174
Direction	without	0.321 / 0.219
	with	0.296 / 0.240

Comparing the tasks where the additional perception cue was taken into account, there is a 3.6 percent decrease in mean error and standard deviation goes down by 1.2 percent.

TABLE 4
Regression Line Parameters Fitted
to Test Results

Constant	Value
a_s	-0.010414
b_s	0.02680727
a_d	0.0261351
b_d	-0.16202086
c_d	-0.14056624
a_σ	0.48538867
b_σ	0.29785412

Pearson's correlation test and observed no significant dependence within any of the parameters.

To create the compensation model we needed a function to describe the trends in the estimation error. For the global scale, we choose a logarithmic function,

$$E_\sigma(\sigma) = a_\sigma \cdot \log(\sigma) + b_\sigma, \quad (3)$$

where σ is the global scale and a and b are constants defined in Table 4. The logarithmic function gave a slightly worse fit compared to a second order polynomial, but since the first order derivative became negative at 6.5 in parameter space, we found the second order unsuitable to describe the trend, since the average estimation error was not decreasing (see Table 2).

To find a function for the effect of direction, we fitted a periodic cosine function,

$$E_d(\alpha) = a_d \cdot \cos(\alpha + b_d) + c_d, \quad (4)$$

where α is the angular difference between the two sets of particles and a , b and c are defined in Table 4. Since the trend in error is cyclic, we found the cosine function to be better suited than a higher order polynomial.

Finally, we fitted a second order polynomial function for the speed multiplier parameter, which we constrained to be zero when the particles were moving with the same speed,

$$E_s(s) = a_s \cdot s^2 + b_s \cdot s - (a_s + b_s), \quad (5)$$

where s is the speed multiplier, and a and b can be looked up in Table 4.

Each function provides an estimated error for the given parameter. The final compensation function should provide a scaling function for the velocity to compensate for the total error from all parameters. In addition, the average error at the base level for each parameter would be contained in each function. This was solved by only including the change in error for the scale parameter and the direction parameter. The compensation function combines the error functions in the following manner:

$$\begin{aligned} C_\sigma(\sigma) &= 1 + E_\sigma(\sigma) + E_\sigma(1) \\ C_d(\alpha) &= 1 + E_d(\alpha) + E_d(0) \\ C_s(s) &= 1 + E_s(s) \\ F_c(\sigma, s, \alpha) &= \frac{1}{C_\sigma \cdot C_d \cdot C_s}, \end{aligned} \quad (6)$$

where F_c is the final compensation function, graphically depicted in Fig. 5.

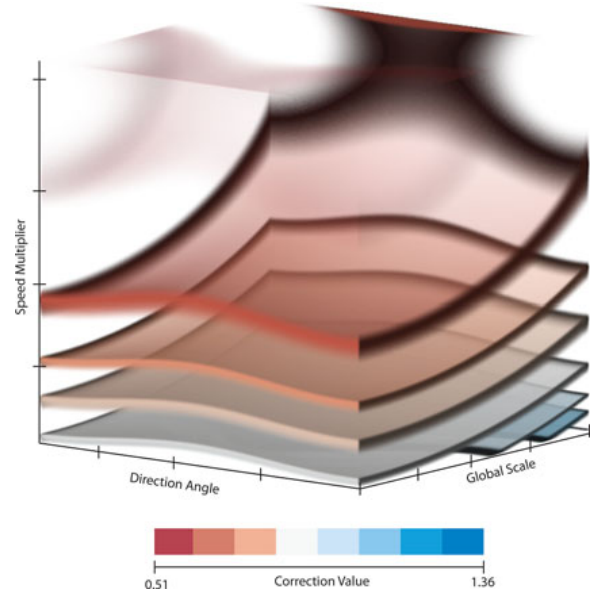


Fig. 5. A volume rendering for the 3D function defining our compensation model.

Before we move on the two following sections where we present our evaluation and adjustment of the compensation function, we first present how our model can be evaluated in relation to the psychophysics literature, in particular to Steven's power law [18].

5.2 Validation against Psychophysics Theories

In our method described above, we take a linear approach in computing the errors between the reference and test displays, and we compute the difference between the multiplication factor we set and the one given by the user. Here, we validate this error computation approach with a comparison to Steven's Power Law theory [18] from psychophysics literature.

As briefly mentioned in Section 2, Steven's theory states that there is a non-linear relationship between perception and stimulus that follows a power law function in the form $y = kx^a$. If we apply this to our case where the stimulus is the velocity of the *test particles* \mathbf{v}_t , this formula turns into $\mathbf{v}_{pr} = k\mathbf{v}_t^a$. Here, \mathbf{v}_{pr} indicates the velocity (of particles) that subjects should be perceiving according to Steven's theory.

Following Steven's theory, we adjust the ground truth while evaluating the s values that subjects provide for each test. This leads to new error values and distributions. Refer to Appendix for further details on how these computations are carried out. We first compute the mean of the error distributions for the new values. The average error for the global scale related tasks is 0.290 when the modified error values are considered. When compared to the previous error distribution, we observe no significant difference. However, for the other tasks (i.e., chroma versus luminance and direction), the average error values are significantly lower, indicating an overall underestimation. This is very likely due to the fact that in these other tasks speed difference is not the only varying stimuli and a more complex psychophysical is needed.

We now focus our attention to check whether we observe similar trends in estimation error when the error values are

TABLE 5
Results of a General ANOVA Test Treating Each of the 10 Bins as a Separate Group with Errors Computed by Considering Steven's Power Law

Parameter	F -value	p -value
Global Scale	3.9193	9.4186×10^{-5}
Direction	0.4316	0.6494
Speed Multiplier	74.9567	4.9585×10^{-121}
Chroma	0.5580	0.8312

computed according to the Power Law model. In order to do that, we perform a one-way ANOVA test following the same procedure used earlier in this section, i.e., the process that led to Table 1. When the new error distributions are used in the ANOVA test, we arrive at the same result with our approach—there is a significant difference within the bins only for the global scale and speed multiplier parameters (see Table 5). This shows that our earlier observations are in line with the power-law-modified computations.

Moreover, we use the new error distributions to fit the same functions we use earlier in this section to perform a further comparison. We compare the *coefficient of determination* scores [31] (i.e., R^2) for both our functions and the new functions fitted to modified error values. We observe that our functions fit better to the data compared to power law functions.

These reported observations demonstrate that our compensation model is in agreement with the previous related studies that fit Steven's Power Law functions to human psychophysical data and thus provides additional support for the validity of our model.

5.3 First Evaluation Study

After we built the initial version of the compensation model, we continued the process by evaluating it through a new round of perceptual study. Our aim at this stage was to assess the changes in the results due to the modifications by the compensation model. We then aimed to improve the model further, as a result of the investigation of the results from the new user study.

In this second round of the study, we made tests to evaluate the three parameters, namely, speed multiplier, direction of motion, and velocity scale. Note that, we left out the contrast parameter at this stage. We randomly created 486 different combinations of these parameters to build the tasks for this round. For each of these parameter combinations, we created two separate types of tasks. In the first task, the final velocity was modified by the compensation model (i.e., experimental group), and, in the second task the final velocity was set without any modification (i.e., control group). This process led to $486 \times 2 = 972$ tasks in total. Similar to the previous round, we used particles with tails for half of the 486 parameter combinations. The tasks were then distributed randomly to 10 subjects, who have not taken part in the first part of the test.

We started analyzing the results by performing an outlier analysis of the results. First, we removed the corresponding tasks from two specific users since their results exhibit conflicting trends when compared to both the 22 users in the

first round and the other eight users in the second round. Additionally, we removed 16 results after an inspection of their Mahalanobis scores. Here, we use two Mahalanobis scores computed using two sets of variables, 1) speed multiplier versus estimation error 2) global scale versus estimation error. The threshold to mark samples as outliers are 8.3 for the first score and 7.8 for the second. After the outlier removal, the remaining set consists of 756 task results.

To evaluate the impact of our compensation model, we treat the experimental and the control group separately (with/without modification). We observe the effect of each parameter on the estimation error separately. In order to do that, we plot the estimation error against the three different parameters for both of the experimental and control group. These plots can be seen in Fig. 6. Moreover, we fitted regression lines to each plot (Table 6) and computed the average estimation error for the two groups of tasks.

For the tasks that are modified with our compensation model, both regression lines highlight a very significant result. We observe that our compensation model manages to flatten out the estimation error trends for both parameters. Specifically, for relative motion the estimation error trend slope goes down from -0.05 to -0.005 , and for velocity scale the slope goes down from 0.07 to 0.005 . This amounts to an approximately 90-93 percent improvement for the trends in estimation error.

Although the correlation between these parameters and the error is removed, the results show that our modifications lead to an overall underestimation of the velocities. This is clearly seen when the average signed estimation error is observed. The average signed error changed from -0.06 to -0.26 . The same observation is also supported by the placement of the regression lines in the second column of Fig. 6, i.e., the regression line is below the x -axis. This

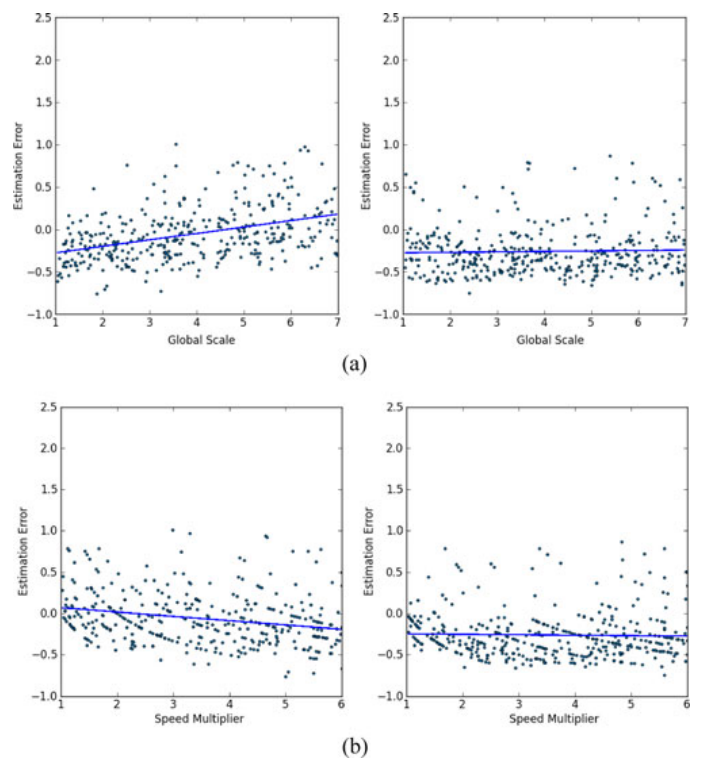


Fig. 6. The regression curves from samples generated in the first evaluation round. The different curves show without (left) and with (right) our compensation model.

TABLE 6
Regression Line Parameters Fitted to Test Results
with/without Modification

Parameter	Modification	Slope	Intercept
Speed multiplier	without	-0.05	0.120
	with	-0.005	-0.241
Global scale	without	0.076	-0.354
	with	0.005	-0.283

observation leads to a modification of the compensation model. We corrected the overall underestimation by inserting a constant ϵ into Equation (6), where ϵ is equal to the average signed error, i.e., -0.26 .

A second modification to our compensation model relates to the *direction* parameter. In order to check for the correlation between estimation error (both signed and unsigned) and the direction of motion parameter, we first group the task results into 10 bins that correspond to 10 consecutive intervals of the motion direction parameter, i.e., each interval spans $360/10 = 36$. Second, we calculate the correlation between the signed/unsigned error and the direction of motion parameter over these 10 intervals. However, there is no significant correlation in any of these intervals. Therefore, we have left out the direction of motion parameter from our compensation model.

As a result of these two modifications, updated compensation model is formulated with:

$$\begin{aligned}
 C_\sigma(\sigma) &= 1 + E_\sigma(\sigma) + E_\sigma(1) \\
 C_\epsilon &= 1 + \epsilon \\
 C_s(s) &= 1 + E_s(s) \\
 F_c(\sigma, s) &= \frac{1}{C_\sigma \cdot C_d \cdot C_\epsilon}.
 \end{aligned} \tag{7}$$

These modifications to the compensation model called for a second evaluation round to assess the efficiency of the updated model.

5.4 Final Evaluation Study

In the final round, we wanted to investigate whether the correction based on the constant found in the previous round would have the desired effect. In addition, we removed the compensation for direction as this parameter had no significant impact. The setup for the final study was almost identical to the previous study with 972 tasks, two sets with equal parameters leaving 486 with compensation from Equation (7), and 486 without any compensation. The tasks were generated using the same constraints from the previous study. Again, we performed the study with ten *new* participants.

From the results we again removed outliers using the Mahalanobis score. We computed the distance for global scale compared to estimation error, and relative motion compared to the estimation error. The distance threshold was determined by visual inspection resulting in threshold of six for the global scale parameter and ten for the relative motion parameter. In addition, samples from one user was deemed unusable, due to having trends conflicting with the other 39 subjects included from all the results. In total 81 samples were removed, leaving 892 samples for analysis.

TABLE 7
Regression Line Parameters Fitted to Test Results with/without
Modification

Parameter	Modification	Slope	Intercept
Speed multiplier	without	-0.051	-0.068
	with	0.007	-0.153
Global scale	without	0.064	-0.37
	with	0.015	-0.240

The results from the final study are shown in Table 7. The average estimation error for sample without compensation was -11.5 percent. For samples with compensation the average estimation error was -17.8 percent. From the previous round we had an average estimation error of -26 percent. Thus leaving the final improvement to be 8.2 percent higher than the previous round.

The parameter impact showed similar improvement in the last round, as we can see in Table 7. In Fig. 8, we can see the regression curves for the tested parameters. The improvement in slope is most prominent for the speed multiplier parameter, the slope for global scale parameter changed from 0.064 to 0.015 . We then conclude that the effect of both parameters have successfully been reduced by more than 75 percent.

Based on the results from the last study, we present the final compensation model as follows:

$$\begin{aligned}
 E_\sigma(\sigma) &= a_\sigma \cdot \log(\sigma) + b_\sigma \\
 E_s(s) &= a_s s^2 + b_s - (a_s + b_s) \\
 C_\sigma(\sigma) &= 1 + E_\sigma(\sigma) + E_\sigma(1) \\
 C_\epsilon &= 1 + \epsilon \\
 C_s(s) &= 1 + E_s(s) \\
 F_c(\sigma, s) &= \frac{1}{C_\sigma \cdot C_d \cdot C_\epsilon}.
 \end{aligned} \tag{8}$$

Where the constants can be found in Table 8. The function is shown as a color map in Fig. 7.

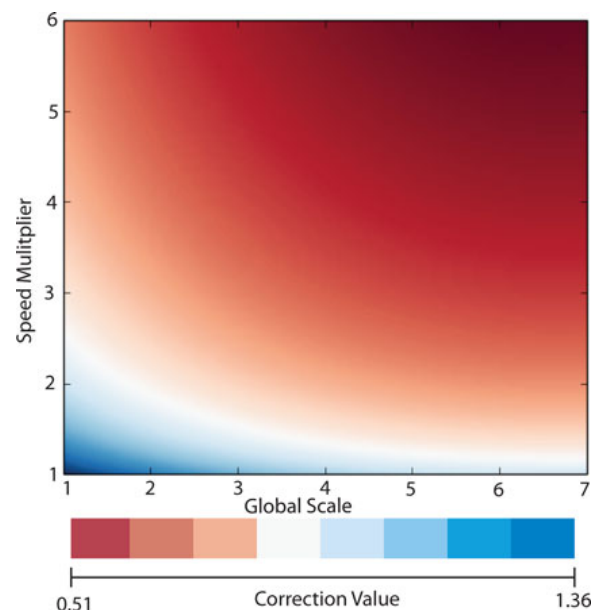


Fig. 7. The final compensation model resulted in a 2D function, depicted here as a height map.

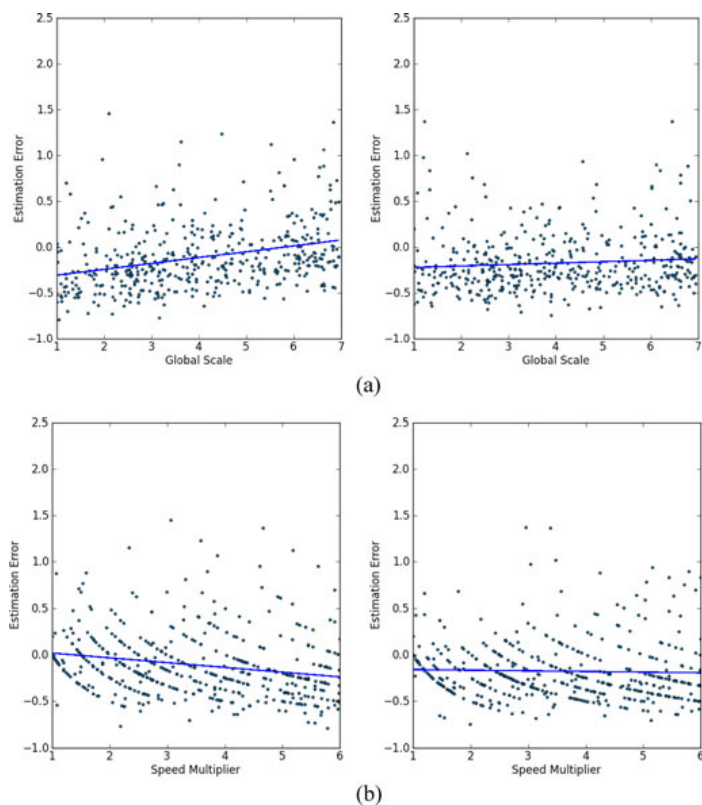


Fig. 8. The results from the final evaluation round. The different curves show without (left) and with (right) the adjusted compensation model.

6 DISCUSSION

Experiments have shown that there was only a systematic change in estimation error when comparing two of the four parameters selected for our study, the speed multiplier and the global scale of the velocities. Increasing the speed multiplier, i.e., when the difference in speed was increased, the perceived difference seemed to scale at a lower rate. Also, when increasing the global scale of the velocities, the estimation error changed from underestimation at low speed (base-speed lower than approximately $1.66^\circ/\text{s}$), to overestimation at higher speeds (above $1.66^\circ/\text{s}$). In this case there seems to be a *sweet spot* where we are most likely to achieve the best estimation, without any compensation. The trend in estimation error compared to speed multiplier remained approximately the same throughout the range of the global scale parameter and should be accounted for.

Although, there was no systematic trend in the signed estimation error, when comparing to the direction of the flow, there was a change in the error magnitude. When

TABLE 8
Constants for Equation (8)

Constant	Value
a_s	-0.010414
b_s	0.02680727
a_σ	0.48538867
b_σ	0.29785412
ϵ	-0.26

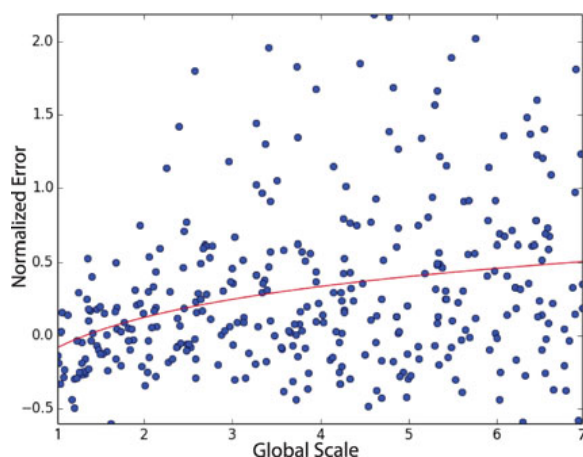


Fig. 9. Error distribution computed following Steven's power law versus the global scale test parameter.

the difference in direction approached 180 degree, there was a small increase in error. However, due to the lack of any systematic change, we were unable to correct for this effect.

From the experiments in previous work, it has been reported a slow-down effect has been reported when gradually changing from luminance contrast to chromatic contrast. In our experiments, this effect was not reproduced. While we should be careful to rule out any impact from contrast-type completely, we might see that this slow down effect is caused not only by different types of contrast, but the gradual change. Mather [14] reported that a change in luminance can create an apparent movement of stationary objects, and it could be that this effect is affecting the perceived motion of moving objects.

From the results using multiple visual cues for velocity encoding, such as comet tails for example, we can only see a slight improvement from simple moving particles. The lack of improvement could indicate that the claim of using multiple visual cues improves the subjects' understanding of speed, is not as prominent as previously believed. However, for higher speeds, comparing tests with a higher global scale of velocities to the general average, there is an improvement of 11.4 percent in estimation error, shown in Table 3. In addition, our results only relate to perceived speed and not to direction, and the additional visual cue provides information not only about the speed, but also about the history of motion as well as that of the particle's current direction. In addition, the usage of comet tails enables the encoding of velocities in still images, which simple particles do not offer, as can be seen in Fig. 2.

To utilize the compensation model in a real world environment, there are certain aspects which should be considered. Since this scales the velocities of particles based on the underlying data, integrating the particle velocities in order to calculate its position becomes an issue. The reason for this is that the actual motion of a particle in the given flow would move at the velocity given by the flow, and arrival time and position would be distorted by the scaling from the compensation model. This can be resolved by scaling the reference frame. Still, the distortion might lead to a less comprehensible visualization. A more proper usage would be to integrate particles for a short time, in order to encode speed at certain position. For longer temporal integration,

other visualization techniques, such as stream lines, would be advised. It is also important to note that visual speed estimation of animated objects will not be very highly accurate. Using moving particles should be used more as an overview, and using a compensated model for depicting the velocities would create a better impression of what exists in the underlying data.

7 SUMMARY AND CONCLUSION

In this paper we have presented a new perceptually based compensation model for using animated particles in visualization. The compensation model is based on the results from a series of perceptual studies, investigating the perceived speed of moving particles. The main goal has been to assess trends in estimation error based on selected parameters. We chose to test for four different parameters, namely global scaling of the particle velocity, the velocity direction, contrast type (iso-luminant versus achromatic), and speed multiplier. Four rounds of studies were performed: A pilot study, an initial study, which was aimed for testing each parameter separately. A second study, which tested the performance of the compensation model. Finally, the last study was conducted to assess the improvements from the previous round.

The results showed that significant trends were only visible in two parameters, global scaling and speed multiplier. A weak trend was found in the direction parameter. Using the trends in estimation error, we constructed a model which can compensate and reduce the effect of each parameter. The improvement was confirmed with a new study where the different parameters were combined. The results from the second study showed a large improvement in the impact factor from the selected parameters. However, the direction parameter was deemed insignificant. In addition, we also found a constant underestimation in speed estimation. Finally, we adjusted our compensation model according to the underestimation constant and performed an evaluation study of the corrected model. This again confirmed the reduction in impact from the significant parameters, as well as it improved the error in estimation compared to the previous study. This work was aimed at 2D flow, and can be used as a starting point for the perception of moving particles in 3D.

The final output from this work is a compensation model for the perceived speed of moving particles. Based on the global scale of the velocities and their relative speed-up factor, we have made an initial step towards a perceptually uniform motion space for animated particles.

APPENDIX

APPLYING STEVEN'S POWER LAW

As discussed in Section 5.2, we support the validity of our approach by modifying the error computations in line with Steven's Power Law [18]. Here, we give details on how we represent the error in velocity estimations in our tests.

Assuming that the human perception of velocity follows a Power Law model, the test subjects perceive the reference

speed v_r according to the following function:

$$v_{pr} = k \cdot v_r^\lambda, \quad (9)$$

where v_{pr} is the perceived velocity. Subjects responded with a estimated speed multiplier between the reference speed and the test speed (i.e., trying to estimate s in Equation (1)). So, the above formula becomes:

$$v_{pr} \cdot s_s = k \cdot (v_r \cdot s)^\lambda, \quad (10)$$

where s_s is the multiplier given by the subject as a response to the test, and s is the true speed multiplier. Since, v_{pr} is not known in Equation (10), we replace it with Equation (9) to get:

$$(k \cdot v_r^\lambda) \cdot s_s = k \cdot (v_r \cdot s)^\lambda. \quad (11)$$

We solve this Power Law model by linearizing this function through taking the *log* of both sides and estimate the k and λ values, in our study these values are found to be $k = 1.4458$ and $\lambda = 0.8428$. Notice that since we try to build a single model (not subject based), we estimate these single values for all the subjects.

The next step here is to modify the s values according to the Power Law model. These values then serve as the expected speed multipliers (i.e., ground truths) in our tests and we denote them as s_{exp} . Assuming a Power Law model version of Equation (1), the computation turns into:

$$v_t = k \cdot (v_r \cdot s_{exp})^\lambda$$

$$s_{exp} = \left(\frac{v_t}{k \cdot v_r^\lambda} \right)^{\frac{1}{\lambda}}.$$

After plugging in the k and λ values and performing the computations we find the estimation errors as

$$s_{err} = s_s - s_{exp}.$$

After the new error distribution is computed, we perform the analysis detailed in Section 5.2. As a companion to the discussions in that section, refer to Fig. 9 that displays the distribution of error values versus the global scale test parameter.

ACKNOWLEDGEMENTS

This work has been carried out within the VERDIKT funded project IllustraSound (# 193170) with support of the MedViz network in Bergen, Norway (PK1760-5897-Project 11). In addition, this work has been supported by Vienna Science and Technology Fund (WWTF) through project VRG11-010 and by EC Marie Curie Career Integration Grant through project PCIG13-GA-2013-618680. The authors would like to express their appreciation for the feedback from experts in visual perception and psychophysics. In particular, they would like to thank Prof. Jan Koenderink, who has encouraged them in the way how they pursued the research for obtaining the motion compensation model. Furthermore, they would like to thank the anonymous reviewers, who have pointed them to relevant psychophysics literature on

perceptual models, LCD and CRT screen particulars, and other suggestions that have helped them how to substantially improve the manuscript.

REFERENCES

- [1] H.-C. Nothdurft, "The role of features in preattentive vision: Comparison of orientation, motion and color cues," *Vis. Res.*, vol. 33, no. 14, pp. 1937–1958, 1993.
- [2] B. Tversky, J. B. Morrison, and M. Betrancourt, "Animation: Can it facilitate?" *Int. J. Human-Comput. Stud.*, vol. 57, no. 4, pp. 247–262, 2002.
- [3] I. Abbott and A. Von Doenhoff, *Theory of Wing Sections: Including a Summary of Airfoil Data*, Series Dover Books on Physics and Chemistry, New York, NY, USA: Dover, 1959.
- [4] D. Acevedo, C. D. Jackson, F. Drury, and D. H. Laidlaw, "Using visual design experts in critique-based evaluation of 2D vector visualization methods," *IEEE Trans. Vis. Comput. Graph.*, vol. 14, no. 4, pp. 877–884, Jul. 2008.
- [5] R. van Pelt, J. O. Bescos, M. Breeuwer, R. E. Clough, M. Eduard Groller, B. T. H. Romeny, and A. Vilanova, "Exploration of 4D MRI blood flow using stylistic visualization," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 6, pp. 1339–1347, Nov./Dec. 2010.
- [6] R. van Pelt, J. Oliván Bescós, M. Breeuwer, R. E. Clough, M. Eduard Grøtiller, B. ter Haar Romeny, and A. Vilanova, "Interactive virtual probing of 4D MRI blood-flow," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 2153–2162, Dec. 2011.
- [7] R. van Pelt, S. Jacobs, B. ter Haar Romeny, and A. Vilanova, "Visualization of 4D blood-flow fields by spatiotemporal hierarchical clustering," *Comput. Graph. Forum*, vol. 31, no. 3, pp. 1065–1074, 2012.
- [8] W. Reichardt, "Autocorrelation, a principle for the evaluation of sensory information by the central nervous system," in *Principles of Sensory Communications*, 1961, pp. 303–317.
- [9] J. P. Frisby and J. V. Stone, *Seeing: The Computational Approach to Biological Vision*, 2nd ed., Cambridge, MA, USA: MIT Press, 2010.
- [10] P. Thompson, "Perceived rate of movement depends on contrast," *Vis. Res.*, vol. 22, no. 3, pp. 377–380, 1982.
- [11] L. S. Stone and P. Thompson, "Human speed perception is contrast dependent," *Vis. Res.*, vol. 32, no. 8, pp. 1535–1549, 1992.
- [12] M. R. Blakemore and R. J. Snowden, "The effect of contrast upon perceived speed: A general phenomenon?" *Perception*, vol. 28, no. 3, pp. 33–48, 1999.
- [13] P. Thompson, K. Brooks, and S. T. Hammett, "Speed can go up as well as down at low contrast: Implications for models of motion perception," *Vis. Res.*, vol. 46, nos. 6/7, pp. 782–786, 2006.
- [14] G. Mather, "Luminance change generates apparent movement: Implications for models of directional specificity in the human visual system," *Vis. Res.*, vol. 24, no. 10, pp. 1399–1405, 1984.
- [15] A. Derrington and G. Henning, "Detecting and discriminating the direction of motion of luminance and colour gratings," *Vis. Res.*, vol. 33, nos. 5/6, pp. 799–811, 1993.
- [16] P. Cavanagh, C. W. Tyler, and O. E. Favreau, "Perceived velocity of moving chromatic gratings," *J. Opt. Soc. Amer. A, Optics Image Sci.*, vol. 1, pp. 893–899, 1984.
- [17] D. Weiskopf, "On the role of color in the perception of motion in animated visualizations," in *Proc. Conf. Vis.*, 2004, pp. 305–312.
- [18] S. S. Stevens, "On the psychophysical law," *Psychol. Rev.*, vol. 64, no. 3, pp. 153–181, 1957.
- [19] J. Ryan and J. Zanker, "What determines the perceived speed of dots moving within apertures?" *Exp. Brain Res.*, vol. 141, no. 1, pp. 79–87, 2001.
- [20] J. Zanker, "Does motion perception follow Weber's law?" *Perception*, vol. 24, pp. 363–372, 1995.
- [21] K. A. Turano and S. M. Heidenreich, "Speed discrimination of distal stimuli during smooth pursuit eye motion," *Vis. Res.*, vol. 36, no. 21, pp. 3507–3517, 1996.
- [22] D. Regan, S. J. Hamstra, S. Kaushal, A. Vincent, and R. Gray, "Visual processing of the motion of an object in three dimensions for a stationary or a moving observer," *Perception*, vol. 24, pp. 87–103, 1995.
- [23] S. Grossberg and M. E. Rudd, "Cortical dynamics of visual motion perception: Short-range and long-range apparent motion," *Psychol. Rev.*, vol. 99, no. 1, pp. 78–121, 1992.
- [24] V. Šoltészova, C. Turkay, M. C. Price, and I. Viola, "A perceptual-statistics shading model," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 12, pp. 2265–2274, Dec. 2012.
- [25] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization," *Inf. Vis.*, vol. 10, no. 4, pp. 289–309, 2011.
- [26] B. J. Geesaman and N. Qian, "A novel speed illusion involving expansion and rotation patterns," *Vis. Res.*, vol. 36, no. 20, pp. 3281–3292, 1996.
- [27] S. N. Watamaniuk, N. M. Grzywacz, and A. L. Yuille, "Dependence of speed and direction perception on cinematogram dot density," *Vis. Res.*, vol. 33, nos. 5/6, pp. 849–859, 1993.
- [28] M. McCool and E. Fiume, "Hierarchical poisson disk sampling distributions," in *Proc. Conf. Graph. Interface*, 1992, vol. 92, pp. 94–105.
- [29] P. Wang and D. Nikolic, "An LCD monitor with sufficiently precise timing for research in vision," *Frontiers Human Neurosci.*, vol. 5, no. 85, 2011.
- [30] F. A. Verstraten, "On the ancient history of the direction of the motion aftereffect," *Perception*, vol. 25, no. 10, pp. 1177–1187, 1996.
- [31] N. J. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika*, vol. 78, no. 3, pp. 691–692, 1991.



Åsmund Birkeland received the PhD degree in informatics from the University of Bergen, Norway, in 2013. His research interests include medical visualization, perception, and image processing. He is a member of the IEEE.



Cagatay Turkey received the MSc degree from Sabanci University, Istanbul, Turkey, and the PhD degree from the University of Bergen, Norway. He is a faculty member and lecturer of applied data science at the gi-Centre at the Department of Computer Science at City University London, United Kingdom. His research interests mainly include the tight integration of interactive visualizations, data analysis techniques, and supporting exploratory knowledge and capabilities of experts. He has a special interest in high-dimensional, temporal data from bioinformatics and the biomolecular modelling domain.



Ivan Viola received the Dipl.-Ing and Dr. techn. degrees from the Vienna University of Technology, Austria, in 2002 and 2005, respectively. He is an assistant professor at the Vienna University of Technology, Austria, and an adjunct professor at the University of Bergen, Norway. His research interests include illustrative visualization, which stands for new visual abstraction methods that are easy to interpret by humans from the perceptual and cognitive point of view, and are related to techniques developed in visual communication and arts. He coauthored several scientific works published in international journals and conferences such as *IEEE Transactions on Visualization and Computer Graphics*, *IEEE Vis*, *EG CGF*, and *EG EuroVis*, and acted as a reviewer and IPC member for conferences in the field of computer graphics and visualization. He is member of the IEEE Computer Society, Eurographics, VGTC, and the ACM.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.

