Perception-Augmenting Illumination

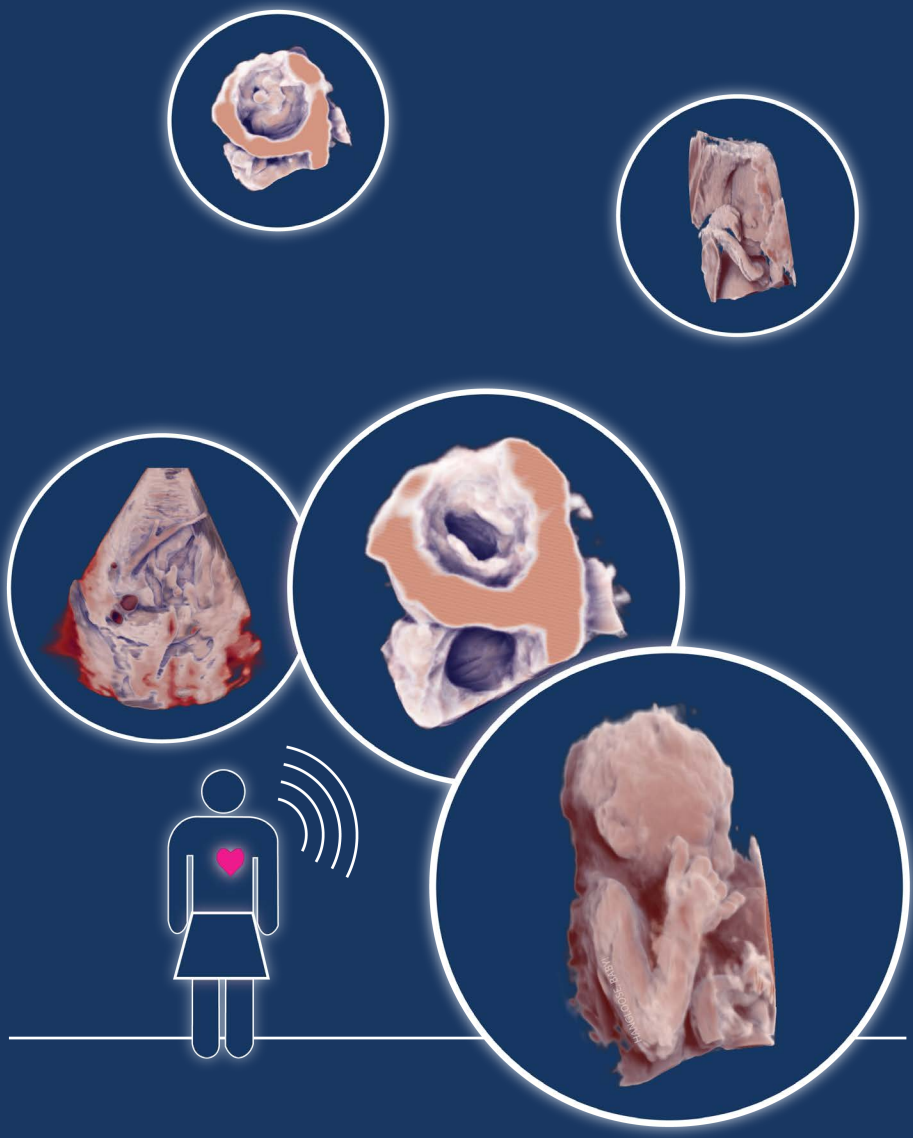
VERONIKA ŠOLTÉSZOVÁ

Dissertation for the degree of Philosophiae Doctor (PhD)

Supervised by Ivan Viola
Co-supervised by Odd Helge Gilja

Department of Informatics
University of Bergen
Norway

9^{th} August 2012



Scientific Environment

The research presented in this thesis has been conducted in the Visualization Group led by Prof. Hauser, at the Department of Informatics, University of Bergen. I have also been enrolled in the ICT Research School at the Department of Informatics at the University of Bergen and the Norwegian Research School in Medical Imaging (MedIm). My research was funded by 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). Furthermore, parts of the research have been carried out in collaboration with the Institute of Marine Research, Norway and Christian Michelsen Research, Norway and Simon Fraser University, Canada, and Munich University of Technology, Germany.

Acknowledgments

First and foremost, I would like to thank Ivan Viola who was my main supervisor and supported me throughout these three years of my PhD. We had man fruitful discussions and also fruitful misunderstandings. He always found time to discuss my progress and somehow managed to squeeze out 200% of the "reasonable" me by saying "when you think you cannot continue anymore, you still have half of your forces left." I am acknowledging the IllustraSound project (# 193170) for financing my PhD research, 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). Parts of my PhD education program were also supported by MedIm – the Norwegian research School in Medical Imaging and the ICT Research School at the Department of Informatics of the University in Bergen.

Thanks also go to people involved in the MedViz network for organizing nice seminars and meetings, and especially to Odd Helge for being my co-supervisor. I am also thanking to all other researchers who were involved in the IllustraSound project for their support: Dag Magne Ulvang, Ola Kristoffer Øye, and Svein Brekke. A special thanks goes to Kari Toverud who was consulting often with me especially in the beginning of my PhD. I am also grateful to all researchers who coauthored my papers for their help and expertise. I am also grateful for being able to have scientific discussions with Meister Eduard Gröller, Prof. Kwan-Liu Ma, Maneesh Agrawala, and finally Timo Ropinski, who made my research stay in Münster, Germany very fun and enjoyable.

I would also like to thank my parents for supporting my education and my activities so that I was able to accomplish my PhD and my sister Barbora for being the "hollow willow" for my complaints during these years. A great thanks goes to my dear Sonička who has been a very good friend to me during these hard times. She is shown in Figure 3.5. Thank you, Soni, for the mysterious cosmetics that you brought for me from the far east that were supposed to prevent my skin from getting wrinkles and dark circles under my eyes. I am wondering why I completely ran out of it now... :)

I am also thanking all the present and former members of the Visualization Group at the Department of Informatics at the University of Bergen: To Helwig Hauser, father of the VisGroup and mother of the CESCG for his advice and discussions; To my office-mate Åsmund Birkeland for his therapeutic skills, to Daniel Patel for being a friend who understands what I mean before I say it and for a great trip to Alaska; To Ovik "Kuvik" Lampe and Endre Lidal for speaking slowly Norwegian; To Július Parulek, Armin Pobitzer and Andrea Brambilla for great skiing trips, where Andrea showed me several possible ways how to break all parts of my body and also for fun we had during our travels; To Július and Johannes "Poking" Kehrer for fun during our travels, especially during the France trip – there were moments I will certainly not forget; To him and Paolo Angelelli, for keeping me company during working weekends and evening. To Çağatay Turkay for his help with statistics and Turkish cooking lessons together with his wife Burcu; To Linn Helljesen for defending my point of view in front of Ivan: that a chair is indeed a place to put a purse.

To Mattia Natali for helping keeping VisTings short and to Jean-Paul Balabanian, after whom I inherited my second monitor and my speakers.

I am of course not forgetting people who I am thankful for my good basis in visualization and computer graphics: again Der Meister, Stefan Bruckner, Michael Wimmer, Werner Purgathofer, Alex Wilkie, and other lecturers from the Institute of Computer Graphics and Algorithms at the Vienna University of Technology and, of course, the CG-Club.

Finally, I would like to say thank you to all my friends, who I have not yet mentioned, and who made my PhD period enjoyable: Burcu, Spiros, Judit, Amani, “umnička” Anna Tikhonova, Thierry, Tetyana and Alex, Satu, Erik and Craig and others.

Abstract

At each stage of the visualization pipeline, the information is impeded by loss or by noise because of imprecise acquisition, storage limitations, and processing. Furthermore, it passes through the complex and not yet well understood pathways in the human visual system and finally to result into a mental image. Due to the noise that impedes the information in the visualization pipeline and the processes in the human visual system, the mental image and the real-world phenomenon do not match. From the aspect of physics, the input of the visual system is confined only to patterns of light. Illumination is therefore essential in 3D visualization for perception of visualized objects.

In this thesis, several advancements for advanced volumetric lighting are presented. First, a novel lighting model that supports interactive light source placement and yields a high-quality soft shadowing effect, is proposed. The light transport is represented by conical functions and approximated with an incremental blurring operation of the opacity buffer during front-to-back slicing of the volume. Furthermore, a new perceptually-founded model for expressing shadows that gives a full control over the appearance of shadows in terms of color and opacity, is presented. Third, a systematic error in perception of surface slant is modeled. This knowledge is then applied to adjust an existing shading model in a manner that compensates for the error in perception. These new visualization methodologies are linked to the knowledge of perceptual psychology and the craft of illustrators, who experimented with visual-presentation techniques for centuries. The new methodologies are showcased on challenging acoustic modalities such as 3D medical ultrasound and sonar imaging.

Related Publications

This thesis is based on the following publications:

(A) **V. Šoltészová**, D. Patel, S. Bruckner and I. Viola. A Multidirectional Occlusion Shading Model for Direct Volume Rendering. In *Computer Graphics Forum (EuroVis 2010)*, 29(3), pages 883–891, 2010 (Level 2)

(B) **V. Šoltészová**, D. Patel and I. Viola. Chromatic Shadows for Improved Perception. In *Proceedings of Non-Photorealistic Rendering and Animation (NPAR 2011)*, pages 105–115, 2011 (Level 1)

(C) **V. Šoltészová**, Ç. Turkay, M. C. Price and I. Viola. A Perceptual-Statistics Shading Model. In *IEEE Transactions on Visualization and Computer Graphics (IEEE SciVis 2012)*, to appear, 2012 (Level 2)

(D) **V. Šoltészová**, R. Patel, H. Hauser and I. Viola. Stylized Volume Visualization of Streamed Sonar Data. In *Proceedings of Spring Conference on Computer Graphics (SCCG 2012)*, pages 13–20, 2012 (Level 1, invited submission of an extended version in Level 2 Computer Graphcis Forum in preparation)

(E) **V. Šoltészová**, L. E. Sævil Helljesen, W. Wein, O. H. Gilja and I. Viola. Lowest-Variance Streamlines for Filtering of 3D Ultrasound. In *Proceedings of Eurographics Workshop on Visual Computing In Biomedicine (EG VCBM 2012)*, accepted, 2012 (Level 1)

The following publication is also related to this thesis:

(I) Å. Birkeland, **V. Šoltészová**, D. Hönigmann, O. H. Gilja, Odd Helge, S. Brekke, T. Ropinski and I. Viola. The Ultrasound Visualization Pipeline. In *Dagstuhl Book*, to appear.

(II) D. Patel, S. Bruckner, **V. Šoltészová**, J. M. Nordbotten. Instant Convolution Shadows for Volumetric Detail Mapping. Submitted to *Transactions on Graphics*.

All manuscripts that constitute this thesis were written during the PhD project and the author of this thesis is also their main author. Ivan Viola, the main supervisor of this thesis, coauthored all manuscripts. Viola contributed not only by his advice and guidance, but also with an inspiration for most of the novel ideas presented in this thesis. **Paper A** and **Paper B** are coauthored by Daniel Patel who helped to maturate the novel ideas during numerous discussions and brainstormings, to collect the related works and to write the manuscript. He also implemented a tool for CIELAB–RGB color conversion. Last but not least, he was a great moral support. **Paper A** is also coauthored by Stefan Bruckner who provided the framework VolumeShop [10], and contributed to the implementation and to the discussions. He also helped revising the manuscript. **Paper C** is

coauthored by Çağatay Turkay and Mark C. Price. Turkay contributed with advice for visual and statistical analysis that was needed for creation of the perceptual model. Price provided help during the experiment-preparation phase and later with the analysis of the collected data. **Paper D** is coauthored by Ruben Patel and Helwig Hauser. Ruben Patel was the collaborating scientists from the Institute of Marine Research who contributed with datasets, his feedback during the design process and with material for the introduction of the manuscript. Hauser contributed with numerous discussions and suggestions to ameliorate the text. The presentation of this paper was awarded by the 2^{nd} best presentation award decided by the audience. The paper itself was also awarded the 2^{nd} best paper award decided by the paper committee and the authors were invited to submit an extended version to the Computer Graphics Forum journal. **Paper E** is also coauthored by Linn Emilie Sævil Helljesen and Odd Helge Gilja who are the clinical partners of the IllustraSound research project. Gilja is also the co-supervisor of this thesis. Helljesen performed the evaluation and Gilja contributed with feedback and amelioration of the text. This paper is also coauthored by Wolfgang Wein who was the collaborating partner and expert on image processing. He provided his advice and contributed to the revision of the manuscript.

All you need is faith and trust... and a little bit of pixie dust

—Peter Pan

Contents

Part I Overview

There is a driving force more powerful than steam, electricity and atomic energy: the will.

—Albert Einstein

CHAPTER

1 Introduction

ONE of the unique features of humans is their highly-developed capability of communication. In other words, they are able to abstract meaning from their sensory input [162], such as hearing or vision. From all sensory modalities, the visual input is the most important one, since as much as 70 percent of the combined input was estimated to be affected by visual sensing [61].

Visualization is a scientific discipline that explores means and methods how to make the visual communication even more effective. The information to-be-communicated is represented in a form of datasets that are then presented to the user in a visual form. The presentation is often not a straight forward process. Even though some datasets do represent real-world phenomena, visualization is not satisfied with a photo-realistic representation of 3D objects. Important is to maximize the information amount received on the side of the user and that is not necessarily achieved by physically-correct visual effects borrowed from computer graphics. Because the communication channel is complex, one cannot assume that the *distal stimulus*, i.e., the real world object from which the information is sampled, equals the *proximal stimulus*, i.e., the sensory input, and the mental image extracted from the information received. For example, a physically correct representation of a real world 3D scene on a computer screen does not mean that the mental image extracted from this visual representation is equal to the real world scene itself.

From the theoretical point of view, at each stage of the visualization pipeline, the information suffers a loss or is impeded by noise. The acquisition process of the real world phenomenon already introduces noise to the signal due to imprecise measuring systems

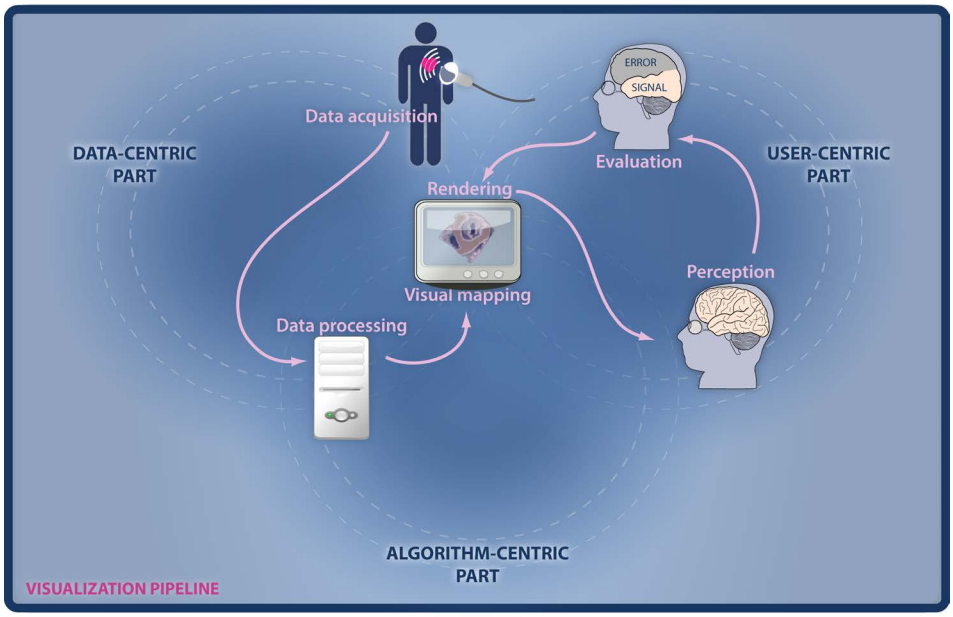


Figure 1.1: The visualization pipeline divided into the data-centric and user-centric parts shown on an example of ultrasound examination.

and storage limitations. Secondly, the data undergoes processing in order to prepare and enhance the data for the presentation stage. The data is modified and this might lead to loss of signal quality. The presentation stage, including data enhancement and rendering, yields a visual stimulus. Furthermore, the information passes through the human visual system (HVS) in which the visual input is processed in a manner that is not yet entirely understood.

We segregated the visualization pipeline into three major groups: the data-centric, the algorithm-centric and the user-centric part. In Figure 1.1, we show this segregation on an example of an ultrasound examination. Some of the stages are not entirely resident within one part. The data acquisition naturally belongs to the data-centric part but also involves a user who is performing the scanning. Data processing means that the data is modified using a specific algorithm. Therefore, also this stage belongs into two groups: data-centric and algorithm-centric. Visual mapping and rendering involves all three groups. Only perception and evaluation of perception is entirely on the side of the user. In this figure, we introduce a new connection that links the evaluation of perception and the rendering stage. If the evaluation confirms an error in perception that the rendering technique can be adjusted to compensate for this error and lead to better match between the proximal and distal stimulus.

Regarding individual stages of the pipeline that contribute to the mismatch between the proximal and the distal stimulus, the match will improve if one of the following con-

ditions applies:

- the acquisition hardware improves,
- the processing stage removes noise more effective with less loss of information,
- the rendering method is well-chosen for the type of data and accounts for the perceptual error induced by the HVS itself.

1.1 Scope and Contributions

The HVS has been and remains an active area of research of several disciplines such as physics, mathematics, psychology, neuroscience and computer science. From the aspect of physics, the sensory input is confined to the patterns of light and to their projection on the retina. Lighting is therefore also essential in visualizing shape and depth relations of 3D objects. The central aim of this thesis is to link the knowledge from psychology regarding the scene perception and the craft of illustration to the illumination models in visualization. The individual contributions of this thesis are the following:

- new knowledge about perception of shape and depth in visualization of 3D objects on 2D screens,
- connection between knowledge from psychology and craft of illustration,
- a new link in the visualization pipeline which allows to adapt the rendering method based on measured perceptual error,
- advancement in global illumination of volumes by utilizing the obtained knowledge and lead to better perception of shape and depth,
- application of our new knowledge to augment the perception surfaces using local illumination (gradient-based shading),
- a showcase of the new models on challenging acoustic modalities: medical ultrasound and sonar imaging.

1.2 Thesis Structure

This thesis consists of two parts. The first part summarizes individual contributions and findings and composites them into a big picture. The second part contains individual publications. The latter are provided verbatim, as they were published, only their formatting was adjusted to fit this thesis and their bibliographies were merged.

The first part starts with an overview of works related to perception and illumination in Chapter 2. A more detailed overview of related works is contained in the individual papers in the second part of this thesis. Since we showcase our methods on ultrasound datasets, we add an overview of important works related to processing and visualization of these modalities. The section about related work is followed up by Chapter 3, the central part of this thesis that is dedicated to perception-augmenting illumination. In Chapter 4, we discuss the application of the new knowledge in the form of better visualization techniques for medical ultrasound and sonar imaging. Finally, we conclude the thesis in Chapter 5.

Pigmaei gigantum humeris impositi plusquam ipsi gigantes vident.

—Sir Isaac Newton

CHAPTER

2 Background and Related Work

THIS chapter starts with a short overview of background knowledge concerning perception of shape and depth inferred from illumination followed by a survey of evaluation methods of aforementioned cues from 2D renderings. Further on, selected approaches of light-transfer computation in volumes are reviewed. For a more detailed overview, it is suggested to study the dedicated sections of individual papers in Part II of this thesis. Since our methods are showcased for usage in other domains, notably medical ultrasound, a selection of related works on processing and rendering of ultrasound datasets is also provided.

2.1 Background on Perception of Depth and Shape

Evolution fine-tuned the HVS to deduce the shape of objects end their spatial arrangement by observing their physical interaction with light. At the first glance, the HVS seems to be constituted by a pathway that goes from the eyes to the thalamus and from there sequentially to the first, i.e., cortical, and then to higher visual areas. This is, however, only on a gross scale. On a finer scale, the pathway is segregated into two major subchannels [83].

Visual Pathways and Sensing of Depth

The subdivision begins in the retina in the arborization of the ganglion cells, i.e., the neuron-type cells that receive input from photoreceptors via two intermediate layers of neurons. The evolutionary older group of ganglions, which is developed also in other

mammals, has larger arborization and thus collects input from more photoreceptors. This type of ganglions is more concentrated in peripheral areas of the retina. They are responsible for sensing of, e.g., motion, spatial position and arrangement, and depth. Livingstone [83] refers to this pathway as to the "Where" system. The other, the "What" system, begins in ganglion cells with smaller arborizations and humans have it in common only with other primates. These ganglions are more concentrated in the foveal area of the retina and give us the ability of recognizing objects, based on their color and details. According to Livingstone, the Where system is colorblind and has high sensitivity to contrast while the What system responds to color but has low contrast sensitivity. Since the depth is processed by the Where system, luminance gradients are indeed sufficient to evoke a sense of depth while chromatic contrast does not create this sense at all.

There is, however, a known optical illusion called chromostereopsis [2, 35] translated as "stereoscopy from color" that seems to be in contradiction with the statements of Livingstone; Two stimuli (red and green or red and blue) located on the same depth plane are perceived to be located at different depth planes. This effect is caused by the transversal chromatic aberration (TCA) of the lens [39]. The refraction index of longer wavelengths (reds) is greater than the index of shorter wavelengths (greens or blues). Therefore, the binocular disparity of a red object is greater than of a blue or a green object, so red objects appear closer. The TCA occurs in the lens, before the light enters the retina, where Livingstone claims the segregation of visual pathways begins. Chromostereoscopy occurs before the start of the visual pathway discussed by Livingstone, so we find her explanations complementing and not contradictory.

Notion and Perception of Shape

The shape of an object is defined by its contour and its surface and is invariant to similarity transformations [120]. The deduction of the object shape involves the low and the high cognitive level. At the low level, the HVS process the intensity variation of shading, deformation of texture patterns, edges and vertices. At the high level, it handles salient features, e.g., occlusion contours that separate the object from the background [147]. Therefore, shading alone cannot convey the complete information about the shape [33] and vice-versa: certain shape cues can be extracted from contours only, but they are more accurately inferred from a shaded image [23].

Since shading is specified by a number of parameters including a direction of the light, viewer position and surface material, the HVS must assume certain parameter settings. For example, the light direction is assumed to be 20°–30° above the viewer [109] and 12° left from the vertical axis [86, 145]. However, this parameter assumption may vary between observers. There is also an evidence that the pictorial relief, i.e., the imaginary relief inferred from a 2D projection of the real world scene of a 3D scene, such as on a computer screen, is systematically distorted [33, 109, 147]. Mamassian and Kersten investigated the perception of local surface orientation [87]. They regarded the slant and tilt of estimated normal vectors of selected points on the surface. By slant and tilt, we understand angles as they are defined in Figure 2.1. Mamassian and Kersten noted that ground truth slants below 20° are generally overestimated and above 20° underestimated. O'Shea and others [109] approximated this distortion by an exponential function. This choice was, however, not a scientifically-grounded but only consistent with the general

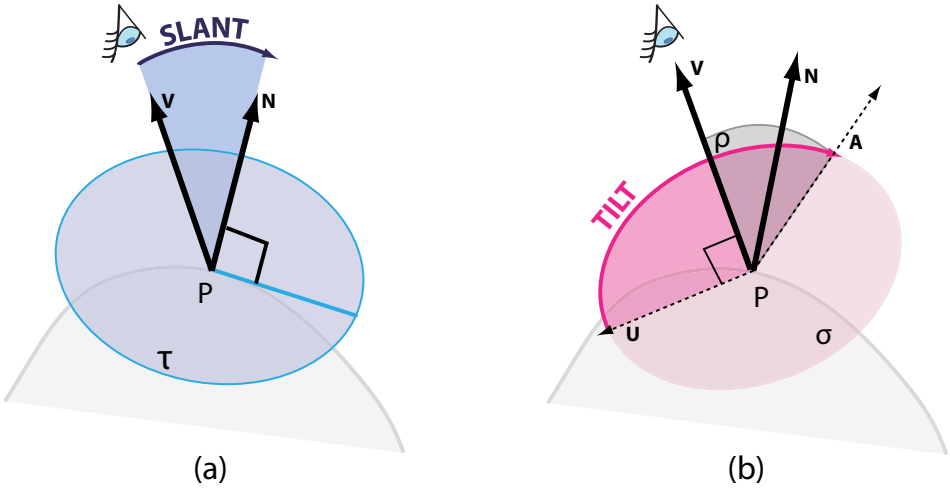


Figure 2.1: A slant angle θ is defined as the angle between the surface normal **N** at a point P and the viewing vector **V**. τ denotes the tangent plane at P and **U** the up vector of the viewer's coordinate-system. σ is a plane such that $P \in \sigma$ and $\mathbf{V} \perp \sigma$ and ρ denotes the plane defined by **V** and **N**. The tilt angle ϕ is then defined as the angle in the left-handed system between **U** and $\mathbf{A} = \rho \cap \sigma$ in the halfplane (ρ,**V**) defined by **N**.

expectation.

Why is the relief perceived flatter even though correct shading is used? This distortion can be explained as follows. Because the two other important sources of depth cues, i.e., binocular disparity and relative motion, that are both zero, the brain receives the information that the rendering is, in fact, flat [83]. In addition, the presence of a frame, i.e., we see that the computer screen is planar in reference to the environment, evokes a subconscious assumption that the rendering is planar as well [151]. The brain receives cues that are in conflict with the cues from shading. As a result, the mental image is flatter than the distal stimulus.

2.2 Quantitative Evaluation of Shape and Depth Cues

The first psychophysical experiments to assess human perception of 3D shapes were conducted in the 19th century. Unfortunately, we dispose only with little information about these experiments and therefore, the conclusions made upon them should be thus considered with caution [147]. The notions object shape and -depth have a certain overlap in meaning. We can quantitatively assess the local shape of objects by surface curvature or local surface orientation (slant and tilt). Relative depth between two points on a surface is nothing else than the z-gradient of the surface-orientation vector since the z-axis in the image space is aligned with the depth. We found that the following experiments are conducted most frequently to assess shape and depth of objects [82, 73, 147].

Relative depth probe task: In this task, observers are exposed to a shaded surface and have to judge the depth of two points. The points are usually marked with dots of

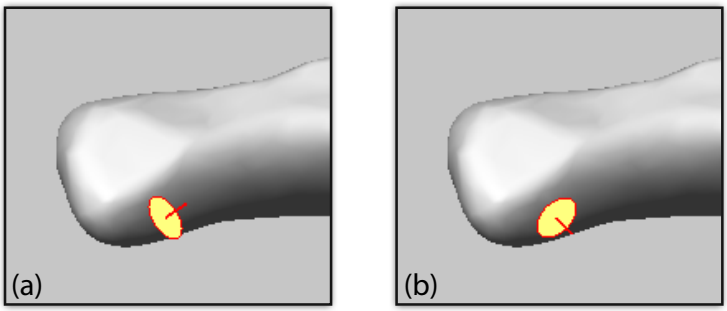


Figure 2.2: Bad (a) and good (b) orientation of a gauge figure.

different colors. The task of the respondent is to select the point which appears to be closer in depth. This task can be varied in such a way that user selects the depth of a selected point in percent between two other reference points.

Gauge-figure task: In this task, introduced by Koenderink et al. [72], a Tissot's indicatrix is employed. A Tissot's indicatrix is a mathematical tool that indicates distortions from projection. It is essentially a circle that together with a perpendicular stick constitutes a gauge figure. On each trial, the respondents are asked to align the indicatrix with the tangent plane of the surface. A bad and a good orientation of a gauge figure is shown in Figure 2.2. The perpendicular stick of the gauge figure should be aligned with the surface normal of the point where the stick intersects the surface. The cues of orientation of the gauge figure come from the perspective deformation of the indicatrix and perspective shortening of the stick. The cues about the surface orientation are inferred from shading.

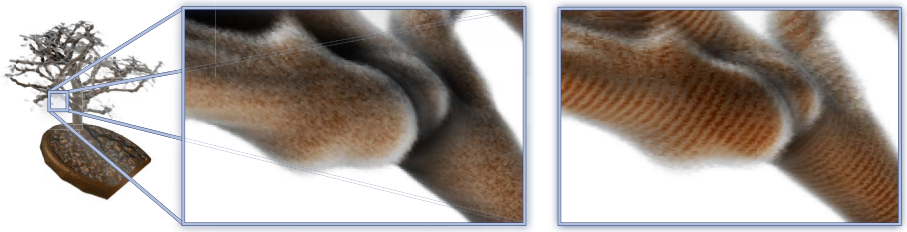
Depth-profile adjustment mask: In this experiment, the surface is displayed overlaid with aligned, equally-spaced dots. On each trial, participants adjust the corresponding dots displayed in another window so that they fit the relief defined by the projections of the dots on the surface displayed in the first window.

Koenderink et al. compared the aforementioned tasks [73] and came to the conclusion that coherent results can be achieved across tasks and even across observers. They also found that the task that is easiest to use is the gauge figure task. The judgment is the most natural and instant, i.e., observers submit their answers based on what they see and not what they deduce from their mental image.

2.3 Illumination

The input of human vision is conceived by patterns of light falling on the retina. Illumination is therefore indispensable in the process of visual sensing and is a *must* in realistic rendering of scenes in computer graphics and 3D volume visualization.

Lighting models vary from simple ones which are easy to evaluate but are less realistic than computationally-demanding ones based on physics [116]. Simple models evaluate only a small neighborhood and are therefore easy to compute. A good representative is the application of the Phong illumination model based on surface normals [117]. According to this model, surfaces are shaded depending on the ambient, diffuse and specular term.



(a) (b)

Figure 2.3: Visualizations of the bonsai dataset: slice-based volume rendering using a view-aligned slice stack 2.3a and using a half-angle-aligned slice stack 2.3b.

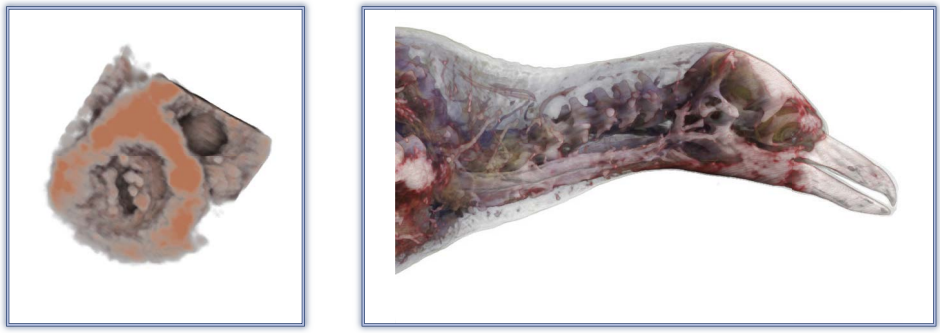
The ambient term is constant, the diffuse term depends on the surface slant from the light direction, i.e., the angle between the light vector and surface normal. The specular term depends on the surface slant from the view direction. For volumetric data, Levoy [80] suggested to use gradients to calculate the shading term in a surface-based illumination model.

We refer to models that are based on surface normals or gradients as to derivative approaches since gradients are first-order partial derivatives. In reality, they are approximated by finite differences, which is a numerically sensitive operation. Therefore, they do not suit well to data with a high level of noise such as MRI [57] and ultrasound [139]. This implies that a noise-prefiltering step needs to be undertaken in order to apply these techniques effectively.

In contrast to local, derivative approaches, there are approaches with a non constant ambient term that is obtained as the occlusion of a given point by a certain neighborhood. We refer to these approaches as integrative since they integrate light energy over a certain neighborhood. If this neighborhood is limited to a certain size, which is small compared to the size of the entire scene, as it is in case of ambient occlusion, we refer to semi-global approaches. Stewart [142] described vicinity shading which enhanced perception by darkening of crevices. Hernell et al. [58] proposed local piecewise integration to approximate occlusion of a certain point by its neighborhood. Their model was later applied in a medical context – for concurrent visualization of structural and functional MRI [104].

Evaluation of global illumination (GI) of a full scene is computationally expensive. For interactive applications, this can be solved either by preprocessing [138, 126] or by approximation [70, 34]. Sloan and others [138] suggested to precompute the radiance transfer for the scene. Ritschel et al. [126] described an efficient encoding scheme for visibility using spherical harmonics. The encoding scheme allows for real-time lighting of dynamic scenes, e.g., volumes where the opacity transfer is adjusted interactively.

Several successful GI-approximation techniques took advantage of volume slicing to approximate scattering. Kniss et al. [70] introduced half-angle slicing which supported forward scattering and shadowing. The approach had two passes. In the first pass, they rendered the scene from the view of the light source while in the second pass, from the point of view of the camera. In both passes, they used slicing in which the slices are aligned to the axis that is set halfway between the view and the light directions. The ad-



(a) (b)

Figure 2.4: (a) Directional occlusion shading model applied on a 3D cardiac ultrasound dataset [134] and (b) image plane sweep volume illumination, image by courtesy of Sundén and colleagues [146].

vantage of this choice of orientation is that the slices were rendered both from the point of light source and the camera. One obvious disadvantage of this method was that it produced artifacts due to non view-aligned slicing as we show in Figure 2.3. More recently, Schott and colleagues [134] proposed a directional occlusion shading model, a single-pass approach that employed a view-aligned slice stack. The model allowed to simulate forward scattering by incrementally blurring the opacity buffer during slice-based rendering. Since the slice stack was view aligned, the light direction was constraint to the view direction. A visualization of a 3D cardiac ultrasound dataset using the directional occlusion shading model is shown in Figure 2.4a. Sundén et al. [146] filled the missing gap and proposed an efficient scattering approximation for raycasters. Rays cast through pixels were synchronized in lines depending on the light direction in the screen space. An example of a visualization using their technique is displayed in Figure 2.4b.

Lindemann and colleagues [82] juxtaposed selected illumination techniques and compared their effectiveness regarding image comprehension. A more detailed overview is contained in the state-of-the-art report on volumetric illumination by Jönsson et al. [63]. Even though precomputation enables a use of the best-quality illumination effects for real-time rendering, the preprocessing stage is very lengthy. At present, the approximation techniques have evolved to such an extent that the results are very realistic and computed at interactive frame rates with no disadvantage of preprocessing delays [77].

Selected works that precompute or approximate on-the-fly realistic light transfer were listed. Some works, however, consider a totally different approach: illustrative shading styles. Artwork inspired Amy Gooch and collective [50] to map the shading coefficient to a cool-to-warm color map. As a follow up work, Bruce Gooch and collective [52], they added features including simple color shadows cast on a plane below the visualized object. Ruiz and others [130] applied illustrative cool-to-warm mapping on ambient occlusion.

2.4 Processing and Rendering of Ultrasound Datasets*

* This section is loosely based on the article *The Ultrasound Visualization Pipeline – A Survey* [8] which was co-authored by Veronika Šoltészová

Since this thesis includes demonstration scenarios of its novel contributions on acoustic modalities, notably on 3D ultrasound datasets, an overview of works related to ultrasound volume visualization is provided as well. Ultrasound (US) scanning has growing popularity as a tool for clinical decision making and is nowadays well established in patient management [107, 108]. Comparing to other imaging modalities used in medicine such as CT, MRI, SPECT and PET, ultrasound has many advantages. It has a low risk of affecting the health of the patients and favorable cost. Furthermore, the scanning procedure is also comfortable for the patient since it can be performed bedside using mobile, hand-carried [48] or even pocket-size scanners [45]. Recently, matrix 3D ultrasound probes came to the market and facilitated real-time acquisition of 3D volumes with high temporal resolution. Unfortunately, ultrasound acquisition is not only prone to "shadowing" (ultrasound waves do not penetrate through bone tissue) and "ghosting" because of reverberations, but also contain a high portion of noise and speckle.

Not only all these features make the US images difficult to interpret but also challenge the visualization technology. It requires solutions that are robust to noise and speckle and in addition satisfy the high temporal resolution of data acquisition. Successful presentation of ultrasound datasets is standing on two pillars: data enhancement and rendering. In this chapter, we provide an overview of both: techniques for enhancing the ultrasound datasets by reducing noise and speckle reduction followed by a survey of available rendering techniques for ultrasound. As it is shown in Chapter 4, the discoveries discussed in Chapter 3 are well suitable for ultrasound visualization and thus contribute to the state of the art.

Data enhancement

The speckle pattern is a natural property of ultrasound imaging. The ultrasound images are inferred from the pattern of sound waves that are reflected at tissue interfaces. The reflection can be anisotropically diffuse, depending on the roughness of the tissue. This causes interference of reflected waves which appears as a speckle pattern in the image. In 2D ultrasound images, the speckle pattern has enjoys clinical appreciation since it reveals certain tissue properties. For 3D ultrasound datasets, the presence of speckle is only impeding the visualization quality and therefore it is worth aiming at removing it.

Recent and most frequently used methods for speckle reduction are based on region growing [18], adaptive filtering [14, 36, 135] and anisotropic diffusion [75]. Local statistical analysis also showed to yield good results. Czerwinski and colleagues [28] described an adaptation of the median filter that preserved boundaries and reduced the amount of speckle. They computed medians for a set of lines passing through each data point and kept the largest median as the new value at the respective data point. Coupé et al. [26] adapted the nonlocal means filter [12]. They used statistical distance for weighting instead of the Euclidean distance of the samples applied in regular bilateral filtering.

Rendering

Presentation of ultrasound datasets is challenging due to the nature of acoustic data, notable because of the inherent speckle and noise. Techniques that are normally used for rendering other modalities might not suit visualization of ultrasound datasets. Nelson and

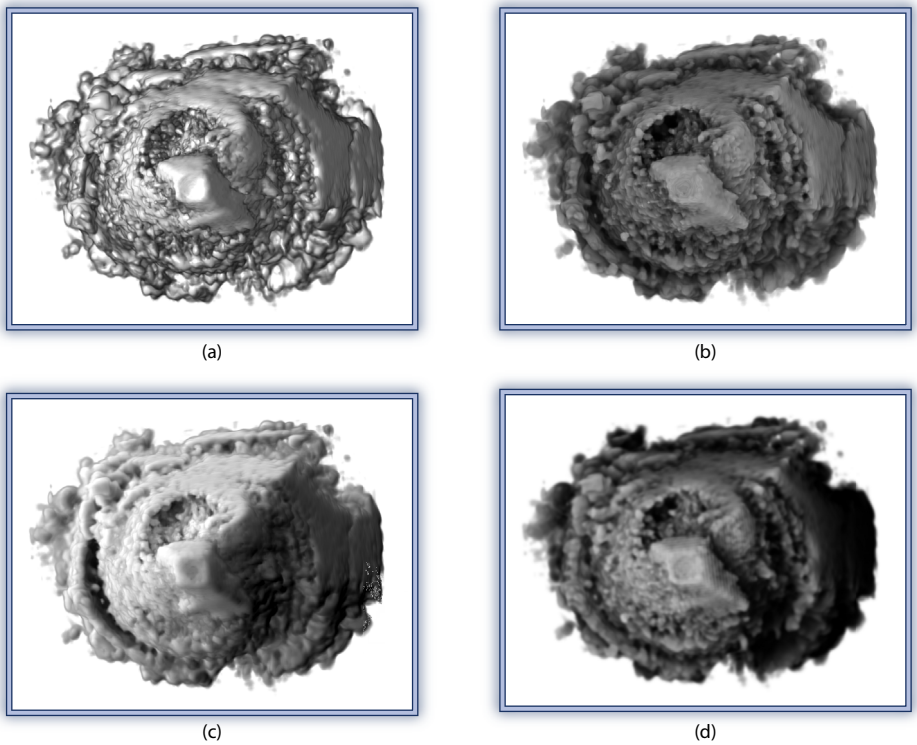


Figure 2.5: Comparison of (a) a gradient-based approach using Phong illumination [117], and gradient-free methods: (b) directional occlusion shading [134], (c) spherical harmonics [81] and (d) the technique of Ropinski and others [127] applied on a 3D cardiac ultrasound dataset. Image by courtesy of Birkeland and colleagues [8].

Elvis [103] examined suitability of then-existing rendering techniques, such as surface fitting and direct volume rendering. Steen and Olstad [141] evaluated selected projections techniques including the maximum intensity projection, average intensity projection, and gradient magnitude projection for fetal ultrasound scans. Provided the good contrast between the fetus and surrounding fluid, the gradient magnitude projection came forth to give the best detail. Careful choice of the opacity transfer function is also crucial. The intensity values in ultrasound datasets to the strength of the echo at tissue interfaces and not to tissue density. Therefore, the design of opacity transfer functions is different than for other modalities. Hönigmann and colleagues [59] proposed a transfer function design for hyperechoic structures located in hypoechoic fluid.

Regarding illumination, gradient-free techniques were indicated to be more suitable for ultrasound visualization [127, 139]. Gradient estimation is an operation sensitive to noise and therefore gradients computed in this way require preprocessing before they can be used for illumination. Birkeland and colleagues [8] compared a selected set of gradient-free to a gradient-based illumination technique (see also Figure 2.5). Some ultrasound workstations offer additional ways for communicating spatial cues, such as depth-based color encoding. This approach is exploiting the effect of chromostereospsis [35].

(a) (b) (c)

Figure 2.6: (a) A workstation screenshot from a 3D cardiac ultrasound examination, and (b) and (c) images of a fetus captured by the HDLive system implemented on new Voluson E8 Expert workstations [44]. Images (b) and (c) are provided by courtesy of GE Healthcare, Healthcare Systems, Ultrasound.

Features close to the viewer are toned with warm and far features with cool colors as in Figure 2.6a. The new generation of ultrasound scanners supports highly advanced material representation of a baby skin and light simulation [44]. Figures 2.6b and 2.6c show and example of prenatal imaging with frontal and rear light source placement.

2

Background and Related Work

Nature and Nature's laws lay hid in night:
God said, 'Let Newton be!' and all was light..

—epitaph of Sir Isaac Newton, A. Pope

CHAPTER

3 Perception-Augmenting Illumination

THE individual contributions claimed in this thesis can be seen as fragments of one story. Papers A, B and C create our theoretical basis for perceptually-augmenting visualization that advances the state of the art at two fronts. First, our new methods for gradient-free illumination inspired by the craft of illustration, enable interactive lighting and expression of shadows. Second, our new model for gradient-based shading, which is founded by our own analysis of perceptual study data, leads to more accurate perception of surface orientation. In Papers D and E, we exploit our theoretical achievements to give new and improved tools for experts of other domains such as visualization of ultrasound in medicine and sonar acoustic imaging in the marine science. In this chapter, we focus on our theoretical achievements. We extract parts from publications collected in Part II that are necessary to understand them as one compact unit.

3.1 Gradient-Free Illumination

As our first step, we investigated the state-of-the-art techniques in illumination of volumes. Our goal was to find a technique that allows for fast computation of global illumination effects. As a result of our search, we selected the directional occlusion shading model [134]. This model approximates high-quality illumination effects such as soft shadowing and forward scattering. Because of these effects, this model is a reasonable choice for visualization of noisy datasets. Furthermore, the approximation is fast and requires no preprocessing. Therefore it is suitable for interactive visualization and even live stream-

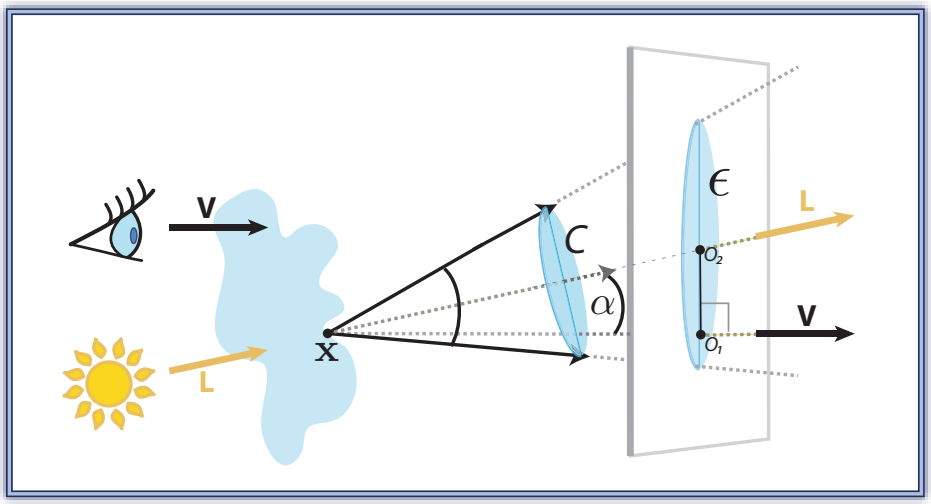


Figure 3.1: A non transparent voxel casts a shadow footprint on a planer receiver. If $\alpha \neq 0$, the footprint is offset by $\mid O_1O_2 \mid \neq 0$ and has an elliptical shape ε.

ing of data. However, consulting visualizations generated with the directional occlusion shading model with a certified medical illustrator revealed an unfortunate limitation of this model.

The light direction is constraint to the view direction but illustrators choose, as a rule, illumination from top-left. We found that their choice of lighting setup is consistent with research in psychology and perception. Humans assume the light source to be 12° left of the vertical axis viewer's vertical axis [86, 145] and 20°–30° above the viewer [109]. With these light settings, the users performed the psychophysics task the best. In order to obtain the best light-possible setting for perception, we proposed the multidirectional occlusion shading model that overcomes the disadvantageous constraint of directional occlusion shading but keeps all its advantages: quality of the shading effect and speed.

3.1.1 Multidirectional Occlusion Shading*

Before we explain how we amended the logic of the directional occlusion model, we clarify why the architecture and notably, the light-simulation approach, constraint the light direction to the viewer as a headlamp. Essentially, every point in the participating, semi-transparent medium acts as an occluder. The semi-transparent participating medium is in our case a volumetric dataset where each voxel intensity is mapped to color and opacity. During the front-to-back slicing process in the viewing direction, each non transparent voxel leaves a shadow footprint on the next slice. The footprints are approximated with Gaussian blurring.

* More details about the multidirectional occlusion shading model are contained in Paper A

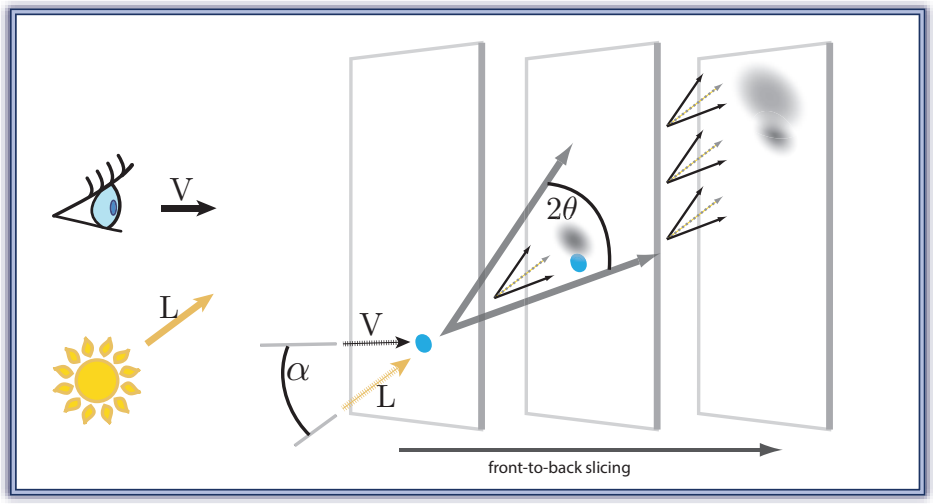


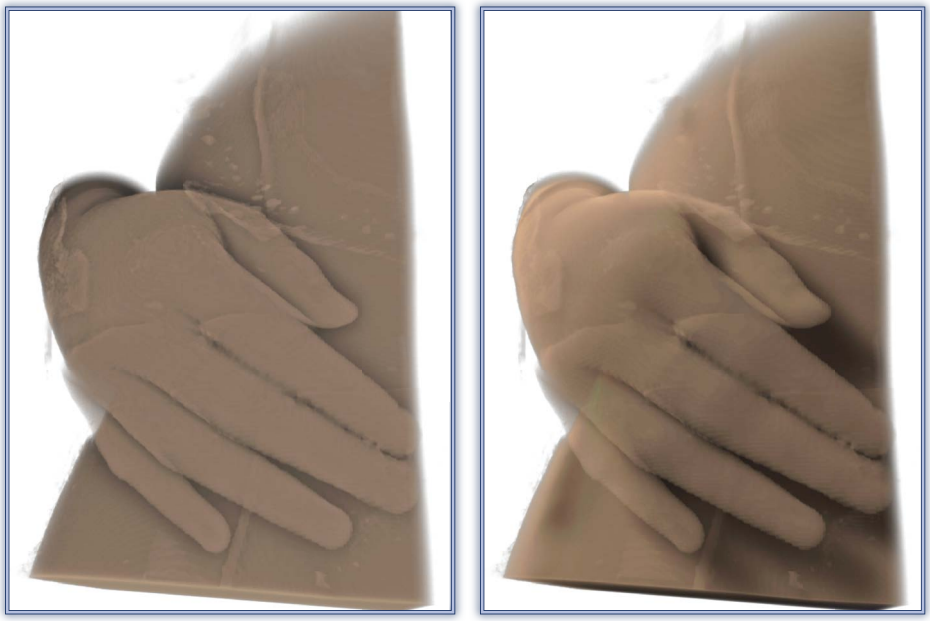
Figure 3.2: Simultaneous front-to-back slicing and shadow propagation. The blue dots indicate not transparent voxels on slicing planes. The underlying shadow buffer is shown in gray scale. The shadow buffer from the previous iteration is blended with new footprints. This process is iterative.

Corrected light propagation scheme

The problem is that this blurring operation assumes the light direction to be aligned with the viewing direction. Light that has a non zero angle with the viewing direction $\alpha \neq 0$ produces footprints transversally-displaced on a planar receiver that are, in addition, centrally asymmetric. We illustrate this situation in Figure 3.1, where a non transparent voxel $\mathbf{x}$ produces an elliptic footprint ε that is displaced by $\vec{O_1O_2}$. If the receiver plane would be perpendicular to the light direction and the light aligned with the viewing direction, the footprints would have zero displacement and circular shape C. In Figure 3.1, we explained shadow casting from a point to the next plane. As we slice the volume, we propagate the shadow through the whole volume as follows. By sampling the shadow buffer produced at slicing step $n-1$, we determine the occlusion of a voxel located on a slicing plane n. The shadow buffer produced at level n is a blending of the opacity footprints of voxels located at slicing plane n and the shadow buffer from the slicing step $n-1$. This process of iterative blurring is illustrated in Figure 3.2.

The blurring kernel

Figures 3.1 and 3.2 show that the blurring kernel encodes the direction of the shadow propagation. If the kernel is a centrally-symmetric Gaussian with no offset, the light direction is consequently parallel to the view direction. According to our definition, the kernel is a function of the angle between the light and the view direction α, the distance between slices and the opening angle of the cone 2θ. The opening angle allows to approximatively simulate the size of the light source. For further detail about the kernel



(a) (b)

Figure 3.3: A human body CT rendered with (a) front lighting that corresponds to the directional occlusion shading and (b) with top left lighting possible with the multidirectional occlusion shading model.

specification, we refer the reader to Paper A.

Results and discussion

Our new concept of iterative blurring allows to interactively select frontal lighting, i.e., light source that is place in the hemisphere defined by the viewer. At the same time, the new model possess all advantages of the previous directional occlusion shading model: no precomputation necessary, a single volume traversal, fast approximation and a high-quality soft shadowing effect. In Figure 3.3, we juxtapose visualizations rendered head-lamp lighting and with top left lighting. This figure illustrates how a lighting setup can support perception of depth. In Figure 3.4, we display another example of the effect of variation of light parameters: direction, slant α and aperture 2θ. In this thesis, we are showing the benefit of multidirectional occlusion shading only on still images, however, an additional advantage of this method are depth cues gained through the interactive modification of the light direction.

We presented our new results to our consulting medical illustrator. She was positive about the new light settings, however, she pointed out that even with the new method we do not have a complete control over the appearance of the shadow. In some areas, the shadowing becomes too strong and hides information. She, as an illustrator, never uses pure black color to express shadows. Instead, she experiments with blue tones, darker

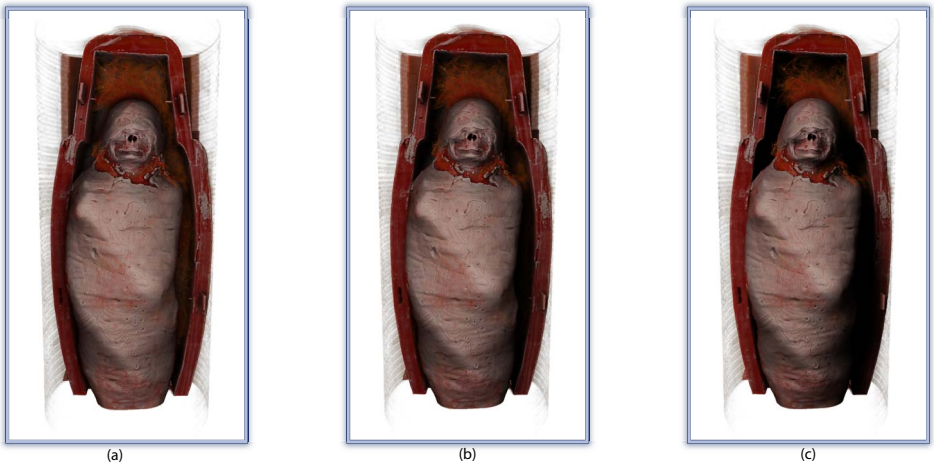


Figure 3.4: A mummy visualized with multidirectional occlusion shading with different lighting parameters: (a) slant angle $\alpha = 0^\circ$ and $\theta = 10^\circ$, which are settings corresponding to front lighting, the only possible setup with directional occlusion shading, (b) from top left and with slant $\alpha = 30^\circ$ as suggested by O'Shea and others [109] and $\theta = 10^\circ$, (c) has the same illumination parameters as (b) except of $\theta = 25^\circ$.

tones of the original color and the contrast colors. Her feedback challenged us to work out a more advanced model that would resolve the information obscuration problem.

3.1.2 Chromatic Shadows*

The craft of illustration has been a good source of inspiration for visualization science so far. To solve the problem of information hiding and to find a better way of shadow expression, we searched for more material in the history of art. Leonardo da Vinci set the rules for shadow colors as follows [111]: "The white, from the sun and the air in the *open air*, has bluish shadows." This is indeed true. On a cloud-free, sunny day, shadows of objects appear to be bluish. As a cause of Rayleigh scattering, the shorter (blue) wavelengths of sunlight are scattered in the atmosphere while other wavelengths pass straight through. Shadows are generated through blockage of direct sunlight, but the ambient light that is scattered from the sky dome is present even in shadows. This light is, indeed, bluish. This effect is clearly noticeable on white material with high albedo, for example on snow, as Figure 3.5 demonstrates both in a photograph and in the artwork [100].

It was the impressionists who started the trend of *open-air painting* that renounced black. Instead, they expressed shadows with blue and complementary tones. As examples, we mention still lifes by Cézanne [16, 17], works of Manet [88], and the *haystack*- and *Rouen cathedral* series of Monet [101, 102]. Figure 3.6 depicts a selection of aforementioned artworks. "I continually search for the blue in shadows": said the post-impressionist van Gogh [111]. Blue shadow tones appear in his works [152], in the works

* More details about the chromatic shadows are contained in Paper B

(a)

(b)

Figure 3.5: Shadows appear in blue tones. This effect is best visible on a clear day on white objects, such as snow: (a) a photograph and (b) an artwork titled *La Pie* (The Magpie) by Claude Monet [100].

Figure 3.6: The Rouen cathedral (a) in full sunlight, harmony of blue and gold and (b) in sunlight from the Monet's *Rouen cathedral* series [102]. *Still life with curtain and flower pitcher* by Cézanne [17] and *The road-menders, Rue de Berne* by Manet [88].

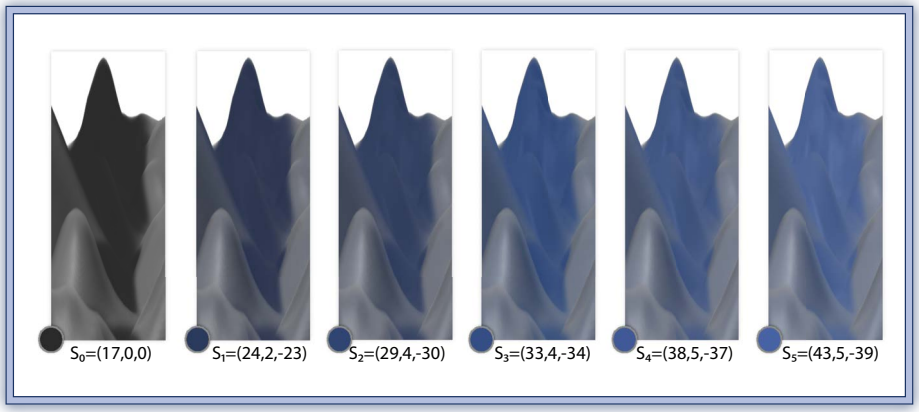


Figure 3.7: A volumetric dataset generated from several sine waves. Using a step transfer function, we visualize a Phong-shaded surface of color O = (60,0,0). We show how the shadow colors $S_0..S_5$ influence the surface perception in the shadow. Shadow colors $S_0..S_5$ have the same perceptual distance from color O. Their luminance contrast is decreasing with increasing chromatic contrast.

of Matisse [91] and in works of contemporary artists as well [24]. Parramón [111] analyzed artistic styles from early prehistoric times until the 20^{th}-century and attempted to formalize the usage of color based on his observation of art history. In his book, he also gives guidelines for expressing shadows that are consistent with the advice of our consulting medical illustrator. He would never use black as a tonal color because it "grays, dirties, and alters tone". He as well segregates the shadow color into three. The main color of shadows is blue. The 2^{nd} color is the local color in a darker tone. The 3^{rd} color is the complementary color of the local color.

The principle

We propose a method that gives the user full control over the appearance of shadows. Soft shadows are in most applications expressed by scaling the color of the shadow receiving object by a scalar factor that we refer to as shadowiness $\mathscr{S}$. This toning is nothing else than a mixing (a linear combination) of the black color RGB = (0,0,0) and the original color of the object. The toning operation results in a variation of luminance between the original color where $\mathscr{S} = 0$ and the total occlusion where $\mathscr{S} = 1$. We propose to express the shadow as a blending of a selected basic shadow color $\mathscr{S}^S_{RGB}$ and the object color $\mathscr{S}^O_{RGB}$. The range of luminance variation in available tones will shrink, but the range of chromatic variation will increase. The local Phong shading, which is dependent on direct lighting and the local slant of the surface from the light direction, is expressed by luminance variation. Since we now express shadows with additional chromatic variation but less luminance variation, we leave more space for the luminance variation from shading. In this way, the observer can see more shading variation and extract better shape cues from the visualization. This effect can be clearly observed in Figure 3.7. We rendered a synthetic object of a uniform gray color CIELAB O = (60,0,0). Color S_0 has no chromatic

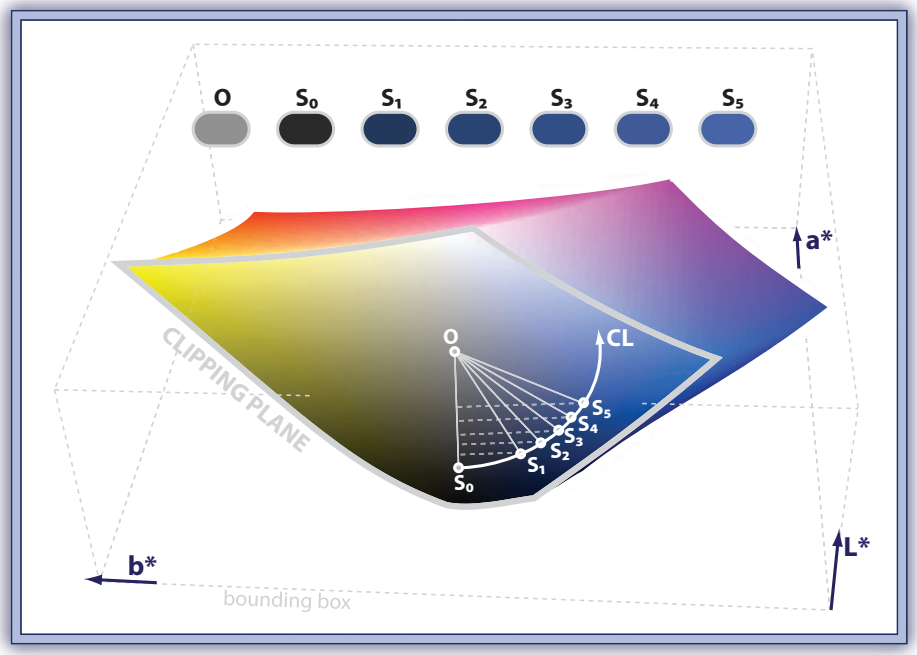


Figure 3.8: Palette of shadow colors $S_0..S_5$ in the CIELAB space: O is the object color and shadow colors $S_0..S_5$ are chosen along the planar arc CL. CL denotes the chromatic-luminance shift and belongs to the circle in the CIELAB color space which is centered in O and has a radius $||OS_0||$. We used the plane define by CL to clip a part of the CIELAB volume to reveal the color plane of our interest. The difference in luminance between S_{n-1} and S_n is constant.

difference to O but colors $S_1..S_5$ have increasing chromatic contrast and decreasing luminance contrast. The perceived contrast of $S_1..S_5$ to the object color O is constant because the Euclidean distance in the CIELAB space is constant. This palette of shadow colors is shown in Figure 3.8. The distance of colors S_n to O is constant, as well as the luminance difference between S_{n-1} and S_n.

The blue tone in shadow and perception

The observation that in open air, the color of shadows is blue, provokes two intriguing questions: Did our vision adapt and develop accordingly? Would there also be some perceptual benefit when using the blue tone in shadows in visualizations? We found no direct evidence, but the following facts indeed support the hypothesis that perception of details in low light condition is better in blue tones. Short-wavelength (blue) cones are the most sensitive of all types of cones [55]. The rods are more than thousand times more sensitive than cones and their sensitivity is shifted toward shorter wavelengths [55].

Another fact is that if we choose any other color as the basic *shadow color* than black, depth perception should be theoretically affected. Why? Any other color has higher lu-

minance than black and therefore, if we choose any other color as the basic shadow color, the available range of luminance variation of available tones will shrink. The variation of luminance is crucial for global depth perception [83]. In this context, blue tones are a better choice than reds and greens. To show this, we converted all color that are reproducible on computer screens as RGB colors, into the perceptually uniform CIELAB color space. In Figure 3.9a, we plotted the L (luminance) values of all basic color tones to show that the luminance of blue tones is inferior to the luminance of all other tones.

Shading enables to detect local features and shadows help to understand the global spatial arrangement. If shading is expressed with a linear combination of the object color *Ob* and black and shadowiness as a linear combination of *Ob* and the shadow color that has non zero chromaticity, a certain range of available colors is obtained. This range (area) depends on the choice of the shadow color. In Figure 3.9b, the areas are compared for blue shadow color (SB), red shadow color (SR), and green shadow color (SG).

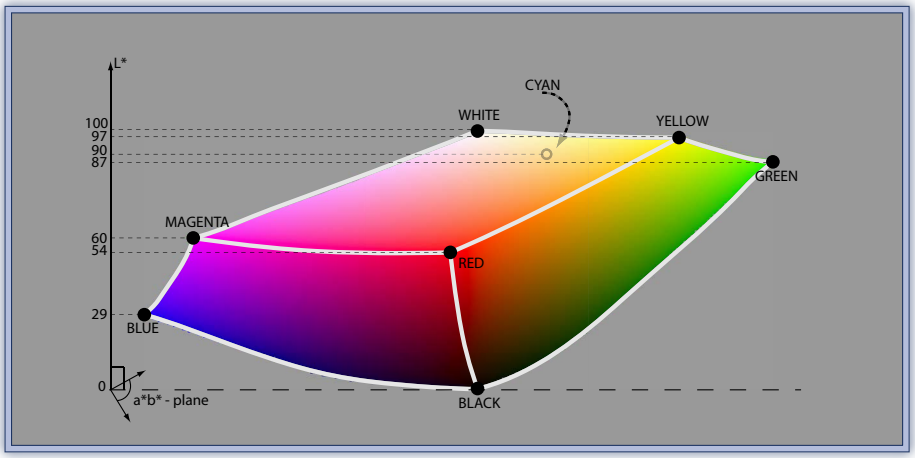
We also performed an initial testing regarding perception of local details and depth. We tested the perception of local details, on synthetic surfaces of constant gray color shaded with local Phong illumination model. We evaluated how precisely users perceive the local surface slant with the gauge figure task [72]. To determine the radius of the just noticeable difference (JND), we presented the participants with the scene visible though a small circular aperture. The task of the users was to increase the radius of the aperture until they are able to perceive color variation within the aperture. In both tests, the gauge figure task and the JND radius task, we found a trend of improvement from S_0 towards S_5. The biggest step towards improvement was from S_0 to S_1.

Finally, we tested the depth perception on volume visualization of objects that are in general difficult to interpret – ultrasound volumes. Users had to relate the depth of a selected point to the depth of two other reference points. We evaluated three conditions: Phong-shaded with no shadows, shadows with a black basic shadow color (luminance only) and shadows with blue basic shadow color (luminance+chroma). The result of the test was consistent with the statement of Livingstone [83]. The depth cues became slightly less accurate for the blue shadow condition and worst for the condition without shadows and with local shading only.

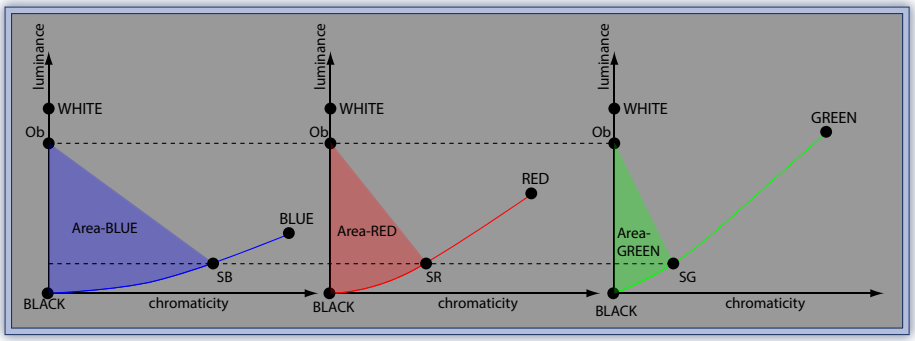
Discussion

With inspiration from illustration craft, we proposed a new concept for expressing shadows and pointed out its perceptual benefits. We investigated the effect of the blue tone for perception of shape and depth. The initial testing allows us to suggest a concept that extracts the best of both, the visible shape cues and sufficient depth cues. For the object color we tested, the biggest improvement in shape perception was observed when increasing the luminance of shadow color only slightly ($S_1 - S_2$), and increase the chromatic contrast accordingly. Even more improvement of shape perception could be achieved if we used S_3, S_4 or S_5, but since we know that decreasing luminance affects depth cues, we suggest to combine $S_1 - S_2$ to have the best of both, shape and depth cues.

We studied the perception only on one example that allowed us to give an initial insight and should encourage further studies that would provide more complete understanding the effects of chromaticity and luminance in shadows on perception.



(a)



(b)

Figure 3.9: (a) An orthogonal projection of the CIELAB color space in other to show the luminance of individual colors. Cyan cannot be seen, but its position is plotted. (b) Available colors if we are using black-toning to express shading and blue, red or green of the same luminance to express shadow: Area-BLUE > Area-RED > Area-GREEN.

3.2 Gradient-Based Shading*

During our previous studies of shape perception, we observed an error trend in user responses. The slant, i.e., the angle between the viewing vector and estimated surface normal) is systematically underestimated. This observation led us to a new concept of the visualization pipeline: we introduced a new link that considers the perceptual evaluation of a rendering method in order to redesign the rendering method itself. The process of evaluation and redesign can be iterative and thus converge towards accurate perception. This concept is illustrated in Figure 3.10.

* More material is included in Paper C

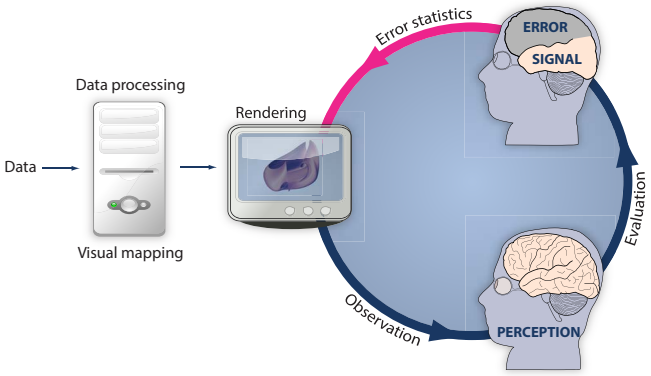


Figure 3.10: The concept of iterative evaluation and design of a rendering technique. The original visualization pipeline contains no cycles and ends at the stage when the image is perceived by the user. The new concept contains a loop; The accuracy of perception is evaluated and the original rendering method is modified based on the measured error in perception.

3.2.1 Analysis of Perceived Surface Slant

As perception literature reports, the surface slant is underestimated [109, 147]. However, there is no model that represents this phenomenon. We obtained our model by analyzing user responses collected and published by Cole et al. [23]. They compared the accuracy of shape cues extracted from line drawings generated by different automatic methods, from line drawings made by hand and from fully shaded renderings. This was a large scale study, the dataset contained ca. 275K solved gauge-figure tasks by 560 participants. From the entire dataset, we focus on perceived slant as a function of the ground truth slant in the fully shaded condition. We fitted a polynomial curve of 4^{th} degree to the data to model the function. However, this representation did not reveal a very interesting property hidden in the data. The systematic underestimation differs for points, where in the ground truth normal points upwards, left, right or downwards. In other words, the underestimation of slant depends on the tilt of the ground truth normal. This finding is displayed in Figure 3.11. We titled the sectors of tilt "north" (normals pointing upwards), "east" (normals pointing right), "south" and "west". The crossing points of the estimation curves and the reference curve are boundaries between over- and underestimation of slant. They are according to Mamassian and Kersten [87] approximately 20° which is, except of the "south" sector, consistent with our finding 15°–25°.

In order to model the human perception of slant, we define two functions $g(\theta_G, \phi_G) = \theta_P$ and $f(\theta_P, \phi_G) = \theta_G$. The mapping g predicts that a surface normal with slant θ_G (G for **g**round truth) will be perceived to have slant θ_P (P for **p**erceived) and f is an inverse function of g. So far, we have only defined four sector curves that map θ_G to θ_P in the interval $[0°, 90°]$. In a similar fashion, we obtained four additional curves for sectors "north-west","north-east","south-east" and "south-west". In order to define g for the whole definition rage of $\phi \in [0, 2\pi]$, we interpolated between the eight curves using Matlab's surface fitting tool [90]. f was defined based on interpolation of inverted sector estimation curves. In Figure 3.12 left, we show a color-coded map of g and $f = g^{-1}$.

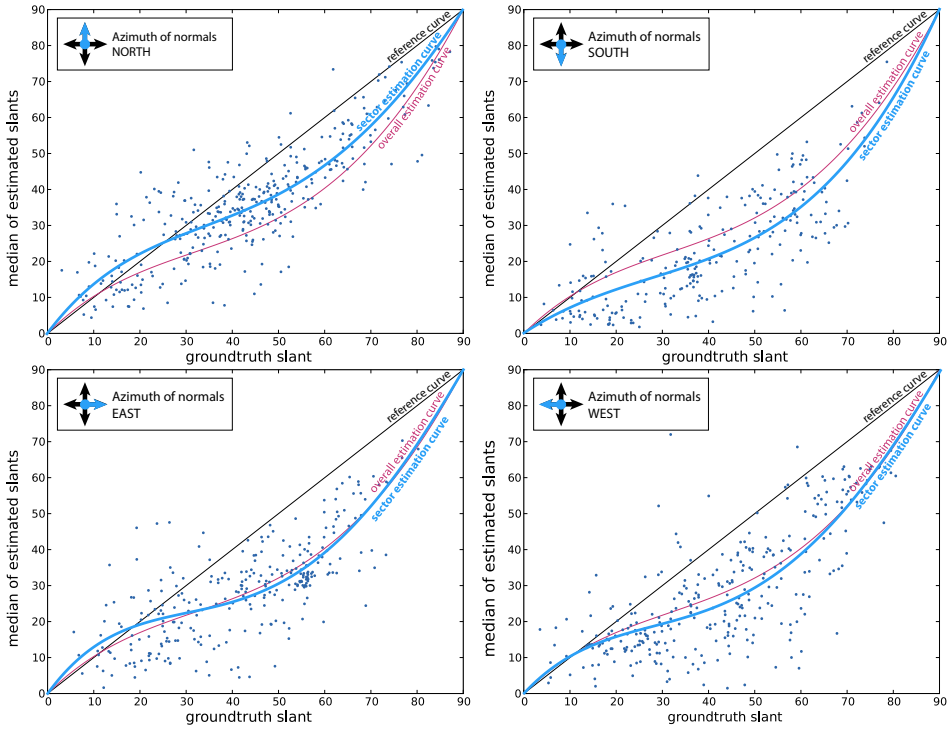


Figure 3.11: Perceived surface slant vs. ground truth slant of fully shaded surfaces extracted from the dataset of Cole et al. [23] for four sectors of tilt. The reference curve indicates a perfectly accurate estimation and the overall estimation curve shows how the dependency would be if we pooled the analysis over all sectors at once.

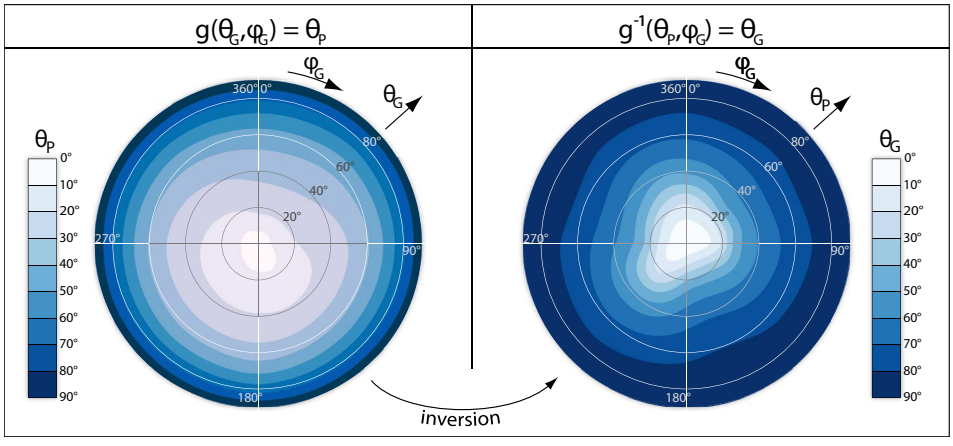


Figure 3.12: Functions $g(\theta_G, \phi_G) = \theta_P$ and $g^{-1}(\theta_P, \phi_G) = \theta_G$ rendered as color-coded plots. Since $f = g^{-1}$, the right plot is also the look-up map which allows to efficiently find the slant angle θ_G of a normal which is perceived as θ_P.

Ideally, such model should be defined for each user, since the bas relief can vary from person to person and for each material, since we know that for example material specularity influences perception of local shape [105, 148].

3.2.2 Perceptual-Statistics Shading

A substantial part of the shape cues extracted from a visualization come from shading of surfaces. We have now concluded that the perception of shape is systematically distorted and fitted a model of this distortion. We aim to exploit this new knowledge to achieve more accurate perception of surface slant.

We want to achieve that a shaded point on the surface is perceived to have a slant equal to its ground truth. A surface normal $n = (x, y, z)$ has ground truth slant θ_G and tilt ϕ_G, both defined in a projective space, but with the original shading, it is perceived to have slant $\theta_P \neq \theta_G$. Therefore, we shade the point on a surface with a modified normal $n' = (x', y', z')$ instead. Function f maps the perceived slants to the ground truth slants. Since we want n' to have perceived slant $\theta'_P = \theta_G$, it should have slant $\theta'_G = f(\theta'_P = \theta_G, \phi_G)$. The new normal n' is now defined by its slant and tilt (θ'_G, ϕ_G). The new normal n' is well-defined by two angles that can be converted to the (x', y', z')-representation. According to our theory, a point shaded with a modified normal be perceived to have the ground truth slant θ_G.

The concept of adjusting normals based on a perceptual-statistics model can be applied to any local surface shading model that causes a systematic error in perception and allows creation of a perceptual model. In Figure 3.13, we show results of the proposed normal (gradient) modification for regular diffuse shading, cool-to-warm shading and also for gradient-based shading used in direct volume rendering. The modification was based on the model we derived from diffuse materials rendered with stochastic raytraycing [22] but ideally, the modification should be extracted for each shading style and each material type in order to achieve the best results.

In order to validate the benefit of the modification based on perceptual statistics, we conducted a new gauge-figure experiment with 40 participants. We compared the accuracy of slant perception with original shading and our shading. We investigated again the original shading condition in order to have a more reliable result. Our model was built on perception of renderings that were generated with a different rendering framework, which could bias our results. We separated the responses according to the shading condition. We analyzed all slants $\in [0°, 90°]$ and also each of the following subintervals separately: [0°,20°], [20°,40°], [40°,60°], and [60°,80°]. The segregation was done on a priori grounds. According to previous [87] and also our own observation, the underestimation is zero for slants approximately 20° and greatest for slants [40°,60°]. Therefore, we predicted different effects in each subinterval. We conducted an ANOVA analysis that confirmed a significant main effect of the subinterval factorization. We observed a trend towards a main effect even when pooled over the whole interval [0°,90°], but this trend failed to reach significance. However, we identified a significant effect of improvement of our shading condition in interval [40°,60°] where the underestimation was highest. In this subinterval, our shading model reached an effect size $r = 0.49$ that is regarded as impressively large within psychological testing literature [20, 21].

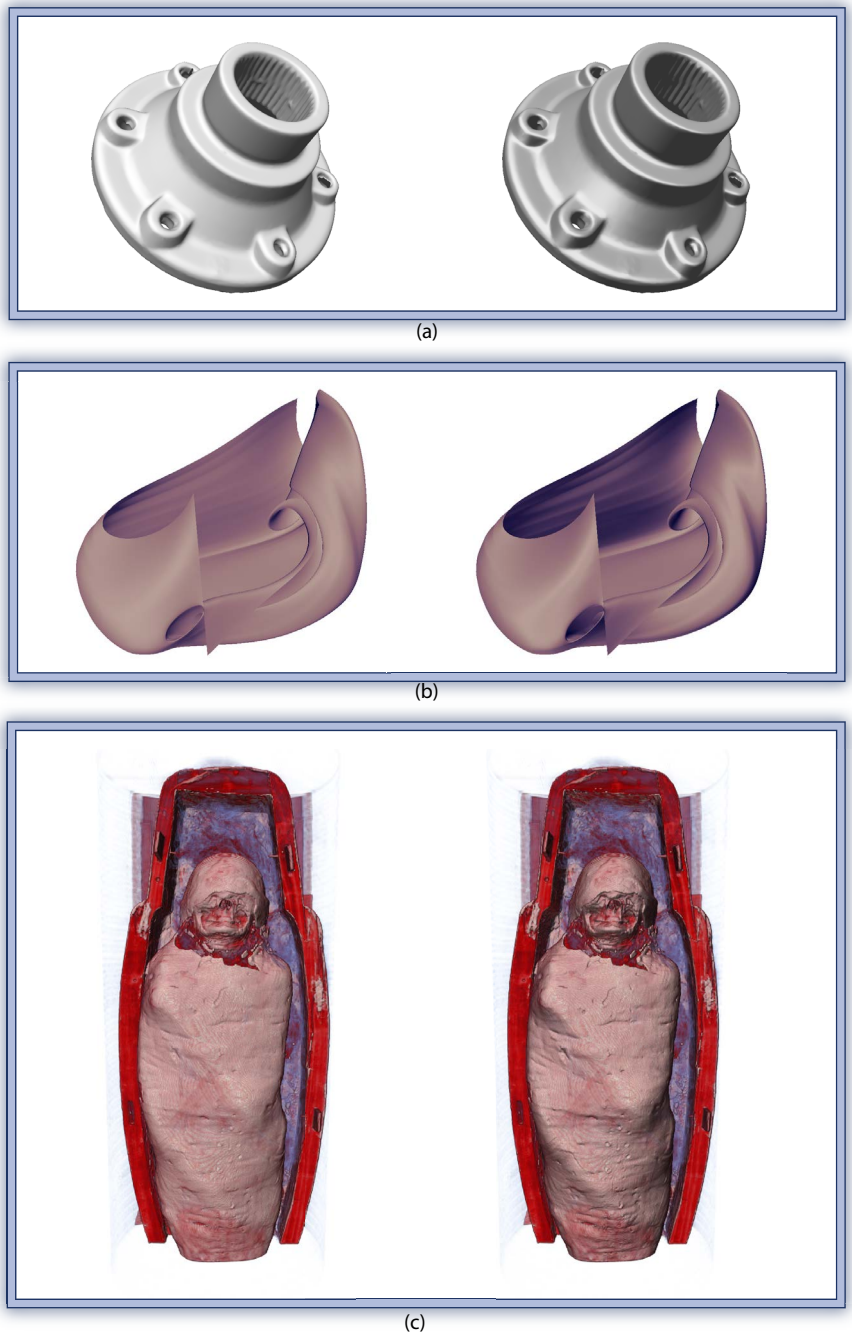


Figure 3.13: Examples of the shading modification applied to (a) diffuse shading of a polygonal model, (b) diffuse cool-to-warm shading of a stream surface, and (c) gradient-based diffuse shading applied in direct volume rendering.

The definition of insanity is continuing to do the same thing over and over, and then expecting different results.

—Albert Einstein

CHAPTER 4

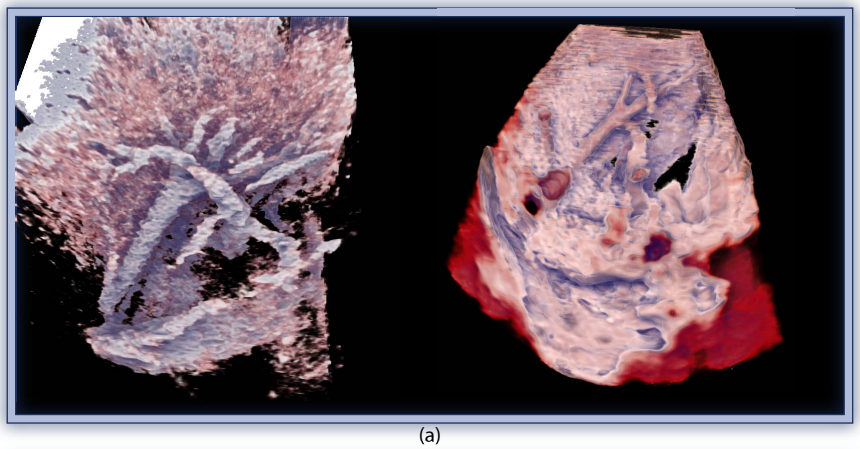
Demonstration Cases

ONE of the goals of this PhD project was to improve visualization of medical ultrasound, but we have extended the impact of our contributions on underwater sonar imaging as well. Both modalities have many characteristics in common. Both of them belong to the category of acoustic imaging modalities. This means that they have similar properties that make the 3D visualization challenging. These properties include speckle and high amount of noise, depth-dependent resolution, shadowing and ghosting due to reverberations. Furthermore, both modalities support live scanning, which presents an additional requirement at the performance of the visualization framework.

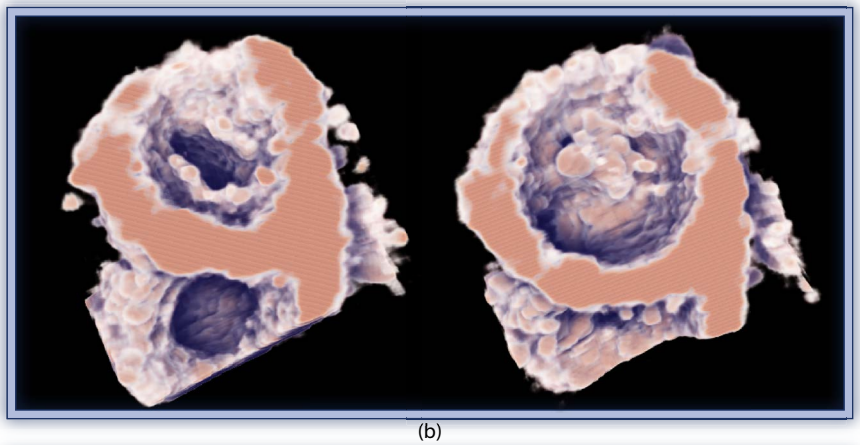
In this chapter, we demonstrate how our theoretical contributions are useful in specific application cases. In Section 4.1, we demonstrate the benefits of gradient-free illumination. Even though gradient-based illumination is not well-suited for the nature of ultrasound, we show a way how to process the dataset so that gradient-based illumination can be used. In Section 4.2, we describe a technology, which we developed in collaboration with marine scientists that is capable of in-situ volume reconstruction and rendering. Amongst other visualization features, it also applies gradient-free lighting.

4.1 Illumination of 3D medical ultrasound

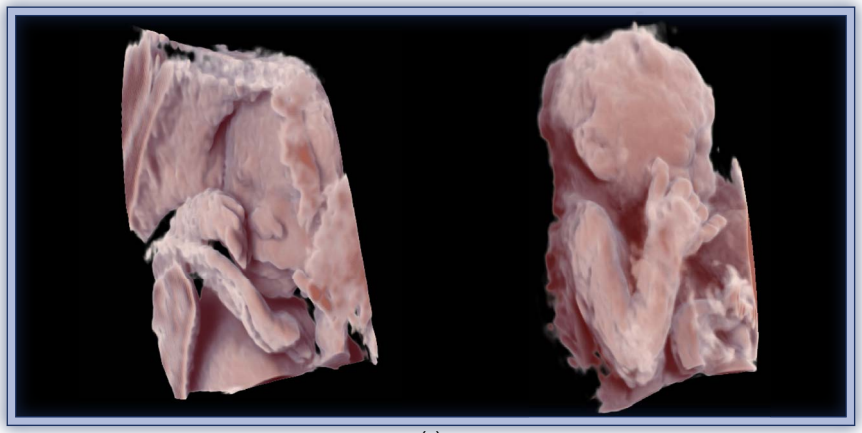
The main advantage of the multidirectional occlusion shading model for our purposes is its interactive performance. A rendering engine that implements this model is then capable of supporting live streaming and in-situ rendering. Furthermore, the model ascer-



(a)



(b)



(c)

Figure 4.1: Visualizations of 3D ultrasound using chromatic shadows: (a) liver vessels, (b) snapshots from a heart cycle, (c) fetuses.

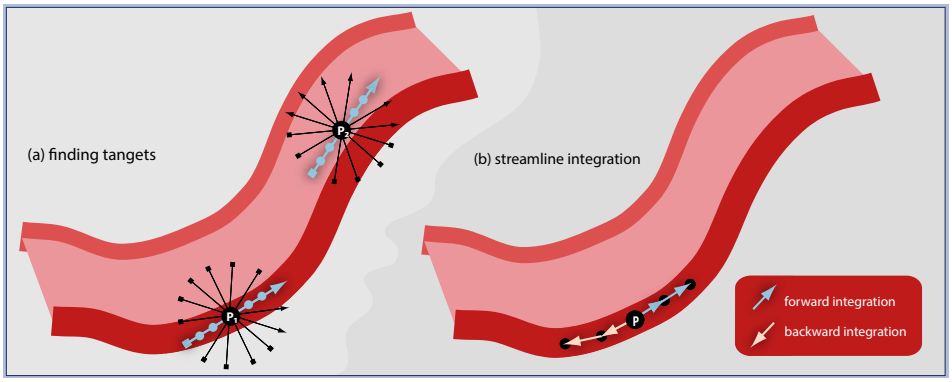


Figure 4.2: Streamline variance filtering. (a) In the first stage, the tangent direction is selected. From all investigated direction, we choose the direction with lowest variance of samples. (b) In the second stage, for each point P, we perform streamline integration seeded in P and following the tangent vector field computed in the first stage.

tains illumination effects of high quality. Finally, the illumination is not based on local properties and therefore does not enhance the presence of speckle and noise. The perceptual improvement of additional chromatic shadows implementation does not consume any additional (noticeable) performance. Therefore, we incorporate it into our visualization framework as well. In Figure 4.1, we display a series of 3D ultrasound visualizations using chromatic shadows.

Application of gradient-based shading on ultrasound data is not straight forward. Local shading based on gradients contributes to local shape cues, but in our case, this visualization style would enhance speckle and high-frequency noise. Therefore, we suggest a preprocessing scheme that effectively eliminates speckle and maintains boundaries. After such a preprocessing, we can also apply the perceptual-statistics shading method that attempts to compensate the inherent flattening of mental image.

Variance-streamline filtering*

In order to decrease the amount of speckle in ultrasound datasets, we propose a filter that locally adapts to tissue boundaries and is robust to noise. There has been attempts to preserve edges by using bilateral filters that apply two types of weighting: distance-based and difference based. This type of filtering can eliminate voxels that have very different intensity values and therefore likely belong to another type of tissue. We are also performing selective averaging but we pursue a different strategy. The filtering operation is divided into two steps that are also explained in Figure 4.2. Even though the steps itself have to be serialized, the computation that is necessary to perform for every voxel in each of both steps can be run in parallel.

1. Determine the tangent direction: The output of this stage is a three-component volume. The vector stored in each voxel is with the highest probability a tangent vector

* More details are contained in Paper E

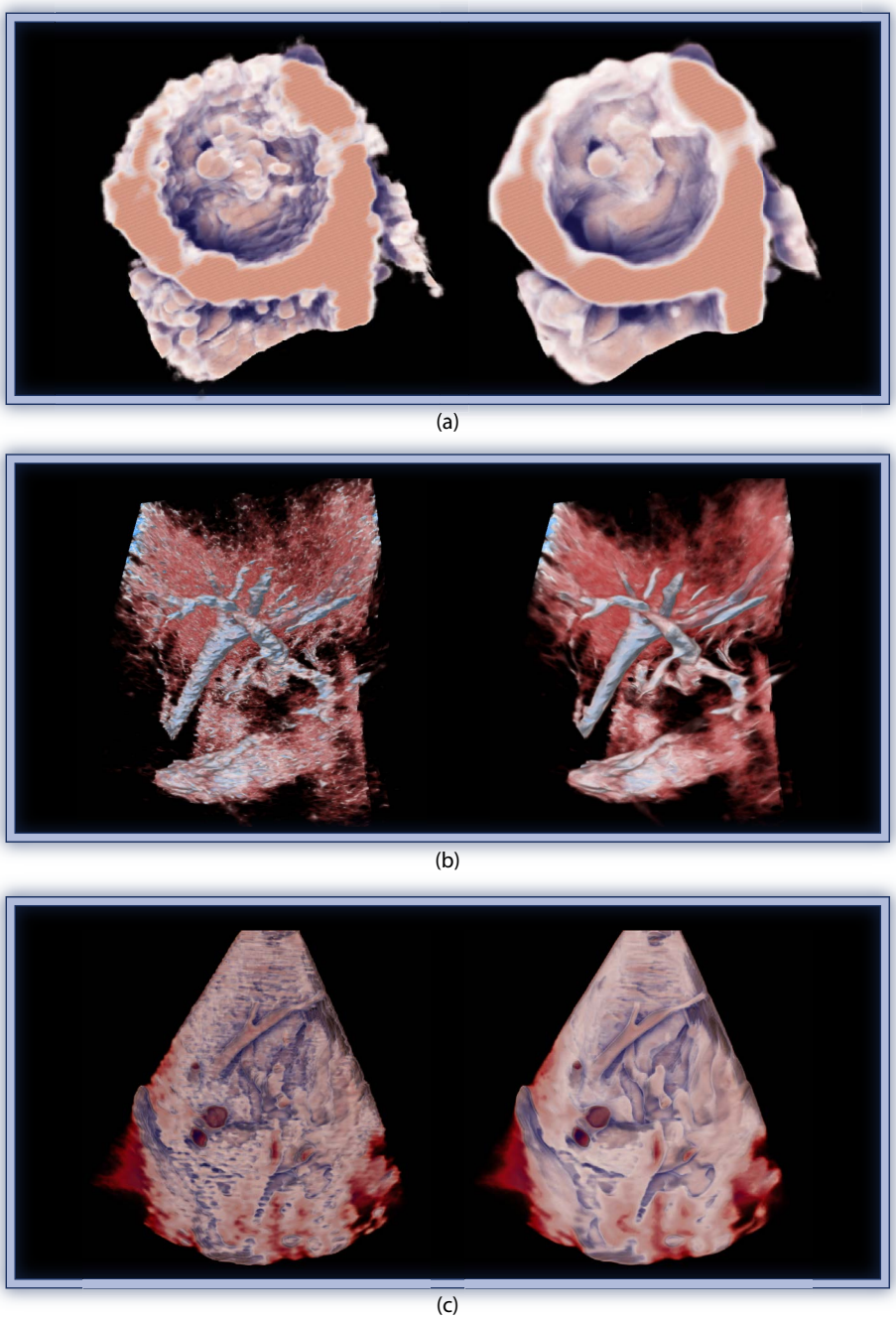


Figure 4.3: Streamline variance filtering applied on ultrasound volumes. (a) human heart visualized with chromatic shadows without (left) and with (filtering). (b) liver vessels with gradient-based shading and no shadowing without (left) and with (right) filtering. (c) liver vessels without (left) and with (right) filtering and visualized using a combination of gradient-based illumination and chromatic shadowing.

to a tissue boundary intersecting the respective voxel. This is performed by running an exhaustive search amongst discretized directions as we show in Figure 4.2a. From all investigated directions, we select the direction that has the lowest variance of samples. In Figure 4.2a, we take three samples forwards and three samples backwards, i.e., $M = 3$. The angular discretization during the search as well as the number of samples $2M + 1$ taken along each direction are parameters that define the accuracy of the selected tangent direction. The parameter M also influences the size of the features that will be filtered. This stage is similar to the principal component analysis of the voxel neighborhood, where the direction corresponding to the smallest eigenvalue is taken.

2. Integrate: In this stage, we filter each voxel with a kernel that is locally adapted to the data. For each voxel P, the operator mask is constructed via streamline integration of the vector field calculated in the previous stage. The integration step consists of forward- and backward integration of N steps. The forward integration starts in P and constructs the "forward" part of the streamline using the vector field that contains the tangent vectors computed in the first stage. The backward integration also starts in P, but construct the "backward" part of the streamline based on the *inverted* vector field. We illustrate this procedure in Figure 4.2b where $N = 2$. N is a parameter that influences the quality of the filtering result.

As a result, we obtain new datasets with decreased content of speckle and high-frequency noise, that are easier to understand, e.g., Figure 4.3a. In addition, the smoothed data allow for computation of better gradients. These can be then used for local illumination, such as our perceptual-statistics shading. Figures 4.3a and 4.3b show visualization that incorporate gradient-based illumination.

In order to assess the quality of this filtering technique, we conducted an evaluation of filtered liver datasets with a gastroenterologist. In liver, the location and the structure of its vessel tree are important. It is especially useful when describing the location of pathologies and also for pre-operative planning. It is crucial that the filtering algorithm should not thin, disconnect or else modify the vessels. To test how well our method preserves tissue boundaries, we compared gastroenterologist's interpretation of the liver vessel tree from visualizations using different filtering methods. In each of the three settings, we rendered a non-filtered and four filtered datasets. We employed three frequently used filtering methods including anisotropic diffusion, Kuwahara filtering, median filtering. The fourth filtering method was our variance-streamline filtering. Inspired by the experiment conducted by Cole and others [22], we asked the gastroenterologist to draw the liver vessel tree over the rendering covered with a transparent foil. We demanded her to filter the speckle mentally, but to draw the vessels according to what she saw. If, for example, a vessel seemed thinned or even disconnected, her task was to draw it thinned or disconnected. Finally, we compared the closeness of her line drawing of filtered datasets to the line drawing of the non filtered dataset. In two out of three settings, the variance-streamline filtering outperformed all other filtering methods. In one setting, the median filtering performed slightly better. In addition to this evaluation, the gastroenterologist made subjective statements concerning the quality of the data in which she consistently preferred variance-streamline filtering.

Our implementation of this method in CUDA exploited parallel computation, but was not yet optimized so that it would fit into a data-streaming framework. A filtering of a dataset 256^3 required approximately five seconds to accomplish on a modern graphics

hardware. Therefore, this method cannot yet satisfy live streaming of volumes at the time being.

4.2 Illumination of sonar data

Another type of acoustic acquisition we showcase our techniques on, is underwater sonar imaging. As we mentioned earlier, acoustic modalities are challenging, especially because of the speckle and noise. In case of sonar imaging, we encountered an additional problem. In order to provide instantaneous volume visualization from 2D sonar images, a reconstruction is necessary and introduces unwanted delays. Direct acquisition of volumes is possible only with 3D sonars. 2D sonars have much lower cost and therefore can be affordable for fishing vessels and regular research vessels as well.

Unfortunately, the visualization technology has stagnated compared to the development of sonar hardware. Currently, the visualization technology supplied with 2D sonars is limited to basic slicing and volume visualization is possible only after a reconstruction step. In-situ volume visualization is possible only with 3D sonar equipment. In collaboration with marine scientists, we developed a new technology that makes in-situ volume visualization, based on streamed sonar images, possible. We proposed a pipeline that coupled the reconstruction and the rendering stage. We based the concept on slice-based volume rendering that allowed us to apply the multidirectional occlusion shading model.

Pipeline

In order to explain how the concept of multidirectional occlusion shading fits into the pipeline, we first give a short overview and briefly explain each stage. The data are acquired using a 2D sonar of type Simrad ME70, which is mounted to sample the water column vertically. In each data collection cycle, referred to as a *ping*, the transducer transmits beams that cover a swat of 140°. Further on, the beams are processed to remove noise and converted to a bitmap. To each bitmap, a timestamp and transformation from a high-end GPS device and an MRU unit are attached. See also stages *acquisition* and *preprocessing* in Figure 4.4.

Previously, the only available in-situ visualization method for 2D sonars was slicing as a direct output of the preprocessing stage. Our system supports instantaneous 3D visualization. As our data input is a stream of images, we construct the volume visualization based on a set of images from the stream. Therefore, the system requires an effective *storage management* unit. The storage unit, also illustrated in Figure 4.4, handles the images, their timestamps and their transformations. In the following *slicer* stage, individual images are transformed to the world coordinate system. The bounding box defined by the transformed images is sliced front-to-back from the direction of the viewer. Each slice of the view-aligned stack is reconstructed in the *slice generator*. The slice generator works on-the-fly during the process of slice-based volume rendering. In order to incorporate the multidirectional occlusion shading model, we add an opacity buffer to store intermediate shadowing information during the traversal of the bounding box.

As we show in Figure 4.5, shadows not only contribute to the realism of the scene, but also significantly improve the depth cues. In Figure 4.5a and 4.5b, we juxtapose renditions without and with the illumination. Shadowing allows to judge distances instantly without

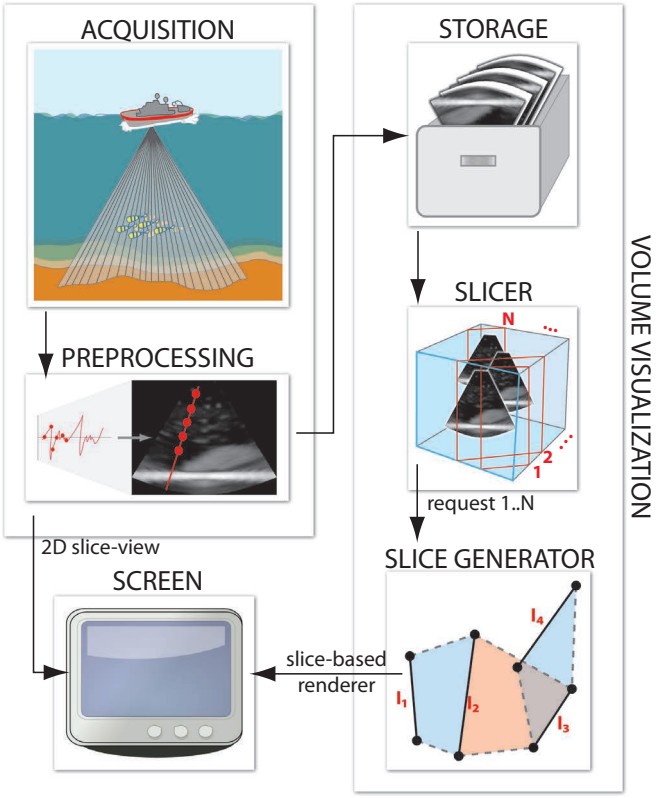


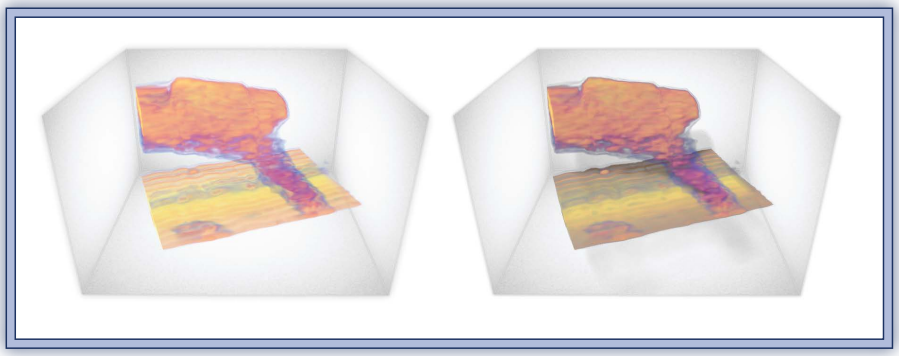
Figure 4.4: Overview of our pipeline.

rotating the scene or other interaction. In Figure 4.5c captures a very large fish school, ca. 100m in length, that was captured in the North Sea in 2010.

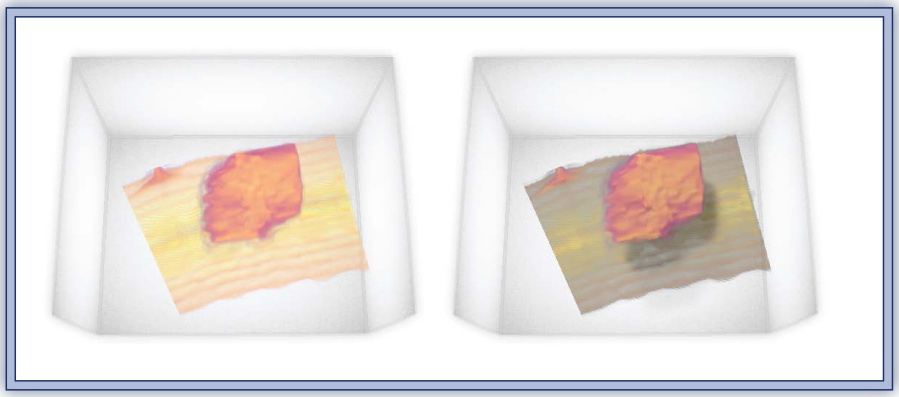
Results

Marine scientists use the sonar equipment to observe and map processes of marine ecosystems. They study the amount and biomass distribution of the stock and also the morphology and behavior within fish schools, i.e., flocks of fish. With limited visualization possibilities, their analysis was mostly limited to quantitative assessment and rudimentary measurements of the fish schools. High quality illumination as well as other features of our tool were new in the visualization technology used in their domain.

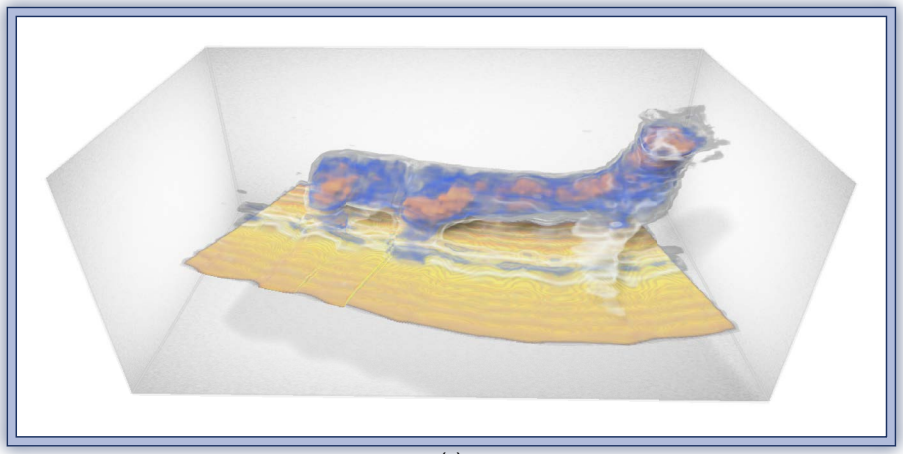
In order to objectively assess the utility of this technology, we contacted independent marine scientists, who were not involved in the design process and a representative of the sonar hardware manufacturer. According to the feedback we obtained, our new technology would bring a clear advantage to the marine domain especially for 3D morphological description of fish schools. The illumination was previously not incorporated in their visualization tools, but was now very well appreciated, since it contributed to a realistic presentation of the scene and better depth perception.



(a)



(b)



(c)

Figure 4.5: Demonstration of multidirectional occlusion shading in the pipeline for volume visualization of streamed sonar images of fish school of sand eel swimming over a sea bottom. In (a) and (b), we juxtapose renditions without (left) and with (right) illumination. (c) is a large fish school of sand eel shown with illumination only.

When I examine myself and my methods of thought, I come to the conclusion that the gift of fantasy has meant more to me than any talent for abstract, positive thinking.

—Albert Einstein

CHAPTER

5 Conclusions and Future Work

THE initial aim of this PhD project was to improve visualization methodologies for medical ultrasound, especially focusing on ultrasound of liver and heart. In the course of the project several inventions were made that aimed to target this modality, such as our research concerning gradient-free illumination. Therefore, the goal of improving visualization of 3D ultrasound has been reached. The application benefit of this thesis is an upper layer that resides on a basis of theoretic advancements, which we would like to point out.

Summary

This thesis advances the state of the art in volume illumination and illustrative visualization. We propose a model that is capable of generating soft shadows in interactive applications, where the user can interactively set the position of the light source anywhere in the frontal hemisphere. To avoid the overdarkening and consequent information hiding in shadow areas, we proposed to express the shadow by variation of chromaticity in addition to variation of luminance. This technique was inspired by the craft of illustration where we observed that artists do not express shadows with pure black color. Since shadowiness is now expressed with an additional variation of chromaticity and less variation of luminance, more effective luminance range is left for expressing shading. Because of this fact, we first proposed a hypothesis that the shape perception will improve with our new shadow specification. We conducted an experiment with users in order to show a trend towards improvement of perception of shape in shadows. Analysis of results obtained in

this experiment led to an interesting observation. The surface slant is systematically underestimated. Even though we found evidence about this effect in literature, there was yet no model that would approximate this feature of human perception of rendered images. We statistically analyzed the error in a dataset from a perceptual study and proposed a model that allows to predict how the surface slant angle will be perceived as a function of the ground truth slant angle. This knowledge allowed us to modify the original illumination approach in order to achieve more accurate perception of surface slant. We conducted a user study that confirmed the improvement.

In order to raise the impact of this thesis from purely theoretical level to the practical level, we showcased several application scenarios of our new illumination methodologies. We showed that our gradient-free illumination methods improve the quality of presentation of challenging acoustic modalities: medical ultrasound and sonar imaging.

Lessons learned

In course of the research performed during this PhD project, we made several invaluable findings that surpass the scope of visualization. Our studies of illustration methodology and perception literature encouraged us to analyze perceptual study data and conduct new experiments. We highlight the most important take-home messages. The first two points are results of literature study and not our own findings. Still, we would like to bring this information to the awareness of the visualization community. All other findings and observations origin from our own work.

- The light direction is assumed to be $20^\circ - 30^\circ$ above the viewer [109] and 12° left from the vertical axis [86, 145]. The default lighting in visualization application should have these settings to maximize the accuracy of shape perception.
- Illustrators' rules for expressing shadows are: the main color is blue, the secondary color is the local color in darker tone, and the tertiary color is the complementary color of the local color [111].
- The additional chromaticity and less luminance variation in shadows allows more perceptually noticeable luminance variance for shading. Therefore, more shape features will be noticeable in shadows.
- Luminance variation is crucial in depth perception [83]. By choosing any tone for the shadow color other than black, the effective luminance variation interval shrinks and the depth cues are affected. More research could be done to correct for this.
- Blue tones have the lowest luminance then other color tones. If a blue tone is chosen as the basic shadow color, we obtain the highest chromatic variation for the lowest sacrifice of luminance variation and less affected depth cues. From this point of view, blue tones are most suitable to express shadows.
- The perception of the surface slant from 2D rendered images is deformed in a systematic fashion. This pattern is almost left-right symmetric. However, the slant of surfaces, which normal vector is pointing upwards is perceived more accurately as the slant of surfaces which normal vectors are pointing downwards. We have not

found a similar discrepancy in slant perception depending on higher order derivatives such as curvature.

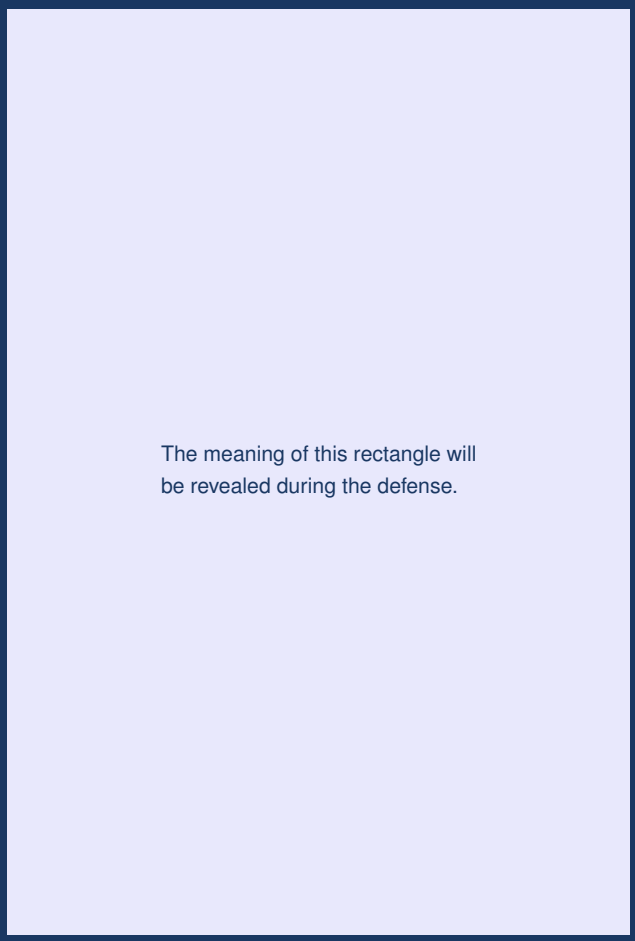
Future Work:

Even though we made several advancements in illumination of volumes and knowledge of perception, this area of research is far from fully explored. As an example, we find worthwhile investigating the perceptual effect of interpolation in CIELAB instead of RGB color space. We would be also delighted to see results of a study that shows the effect other color tones in shadow on perception.

We have noticed that the modification of surface normals changes the appearance of the object material, e.g., shininess. This observation could motivate attempts to characterize a model that adjusts the shading while preserving the appearance of the material. Our shading modification has not reached the optimum, i.e., perfectly accurate perception of slant. This might be due to the fact that shape cues are not solely extracted from shading, but also from other salient features such as contours. The overall perception could be further improved by a global deformation of the object that would compensate for the additional mismatch between the distal and proximal stimuli. This will however deform the overall appearance of the object as a distal stimulus and studies will need to be conducted in order to reveal what happens with the proximal stimulus.

5

Conclusions and Future Work


The meaning of this rectangle will
be revealed during the defense.

Part II Scientific Results

PAPER

A

A Multidirectional Occlusion Shading Model for Direct Volume Rendering*

Veronika Šoltészová[1], Daniel Patel[2], Stefan Bruckner[3], Ivan Viola[1]

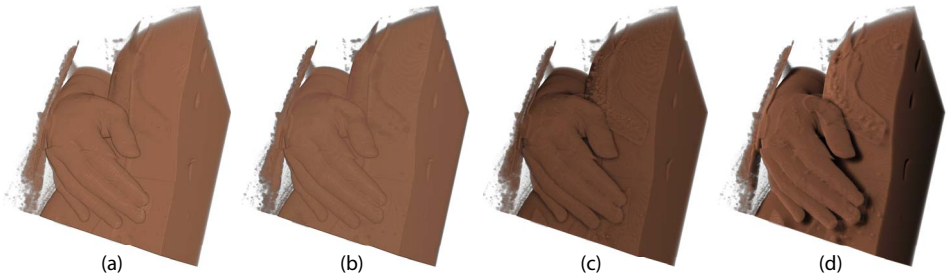


Figure 1: Visualizations of a human hand using raycasting (a), and sliced-based volume rendering (b), both using the Phong illumination. Directional occlusion shading model with a headlamp illumination setup (c) and illumination from top-left (d). Illumination with top-left light position causes that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

Abstract

In 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.

* This article was published in *Computer Graphics Forum*, 29(3):883–891, 2010 and presented at EuroVis 2012 in Bordeaux, France by Veronika Šoltészová

[1] University of Bergen, Norway
[2] Christian Michelsen Research, Norway
[3] Simon Fraser University, Canada

A.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 occlusion shading model introduced by Schott et al. [134] 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 [7, 67, 9]. 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 optimize 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 1a.

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

A.2 Related Work

Levoy [80] 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 provide additional cues. Max [93] gives a comprehensive overview of different optical models for volume rendering.

Yagel et al. [166] employed recursive ray tracing which allows for effects such as specular reflection and shadows. Behrens and Ratering [5] added shadows to texture-based volume rendering by using an additional shadow volume. The model presented by Kniss et al. [70, 71] captures volumetric light attenuation effects including volumetric shadows, phase functions, forward scattering, and chromatic attenuation using half-angle slicing. Hadwiger et al. [54] presented a GPU-accelerated algorithm for computing deep shadow maps for volume rendering. Rezk-Salama [125] proposed a semi-interactive approach for GPU-based Monte Carlo volume raytracing.

Ambient occlusion as described by Zhukov et al. [169] inverts the principle of light-exposure of a point in space to obscurance by its close environment. Dachsbacher et al. [32] refer to obscurance 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 [13, 136]. Knecht [69] and Méndez-Feliu [96] provide comprehensive overviews of rendering techniques based on ambient occlusion and obscurances.

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 [142] 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. [128] 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. [58] used a local approximation of ambient occlusion in volumes to limit computation times. In subsequent work [56, 57], 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. [127] 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. [34] use incremental blurring to achieve shading effects without the use of a gradient. The approach by Bruckner and Gröller [11] 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 [134].

A.3 Multidirectional Occlusion Shading

The **D**irectional **O**cclusion **S**hading model by **M**athias **S**chott et al. (MS-DOS) [134] 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.

A.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 con-

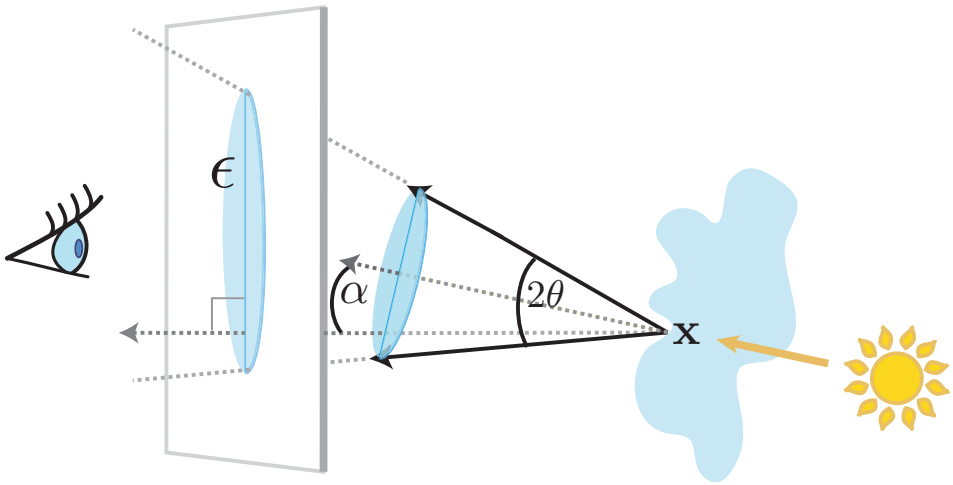


Figure 2: Conical phase function setup: a selected point in space $\mathbf{x}$ scatters light which we approximate by a tilted cone (α = tilt, θ = aperture). The axis of the cone is parallel to the light direction. The projection of the light energy leaves an elliptical footprint ε on a selected viewing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

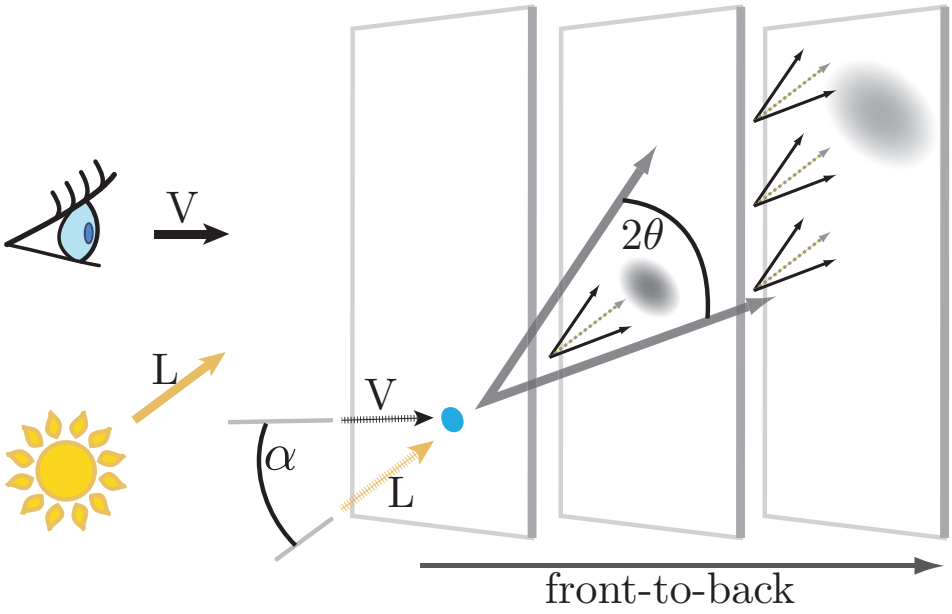


Figure 3: Incremental blurring of the opacity buffer. We use a view-aligned slice-stack composited in the front-to-back order.

sists 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:

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

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 operates on view-aligned slices and assumes that the

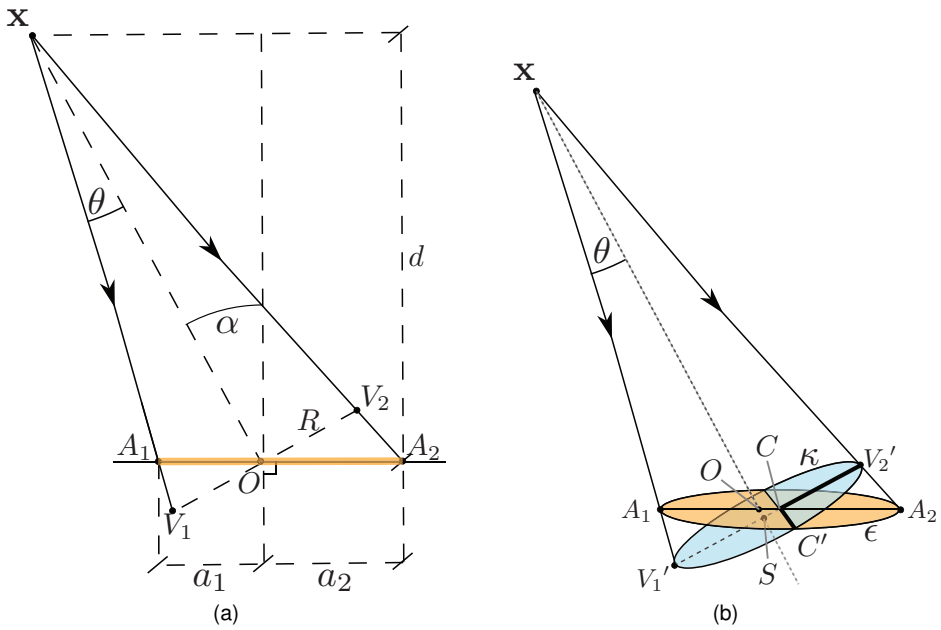


Figure 4: A geometrical description of the cone-shaped phase function: the elliptical cone-section defines a circle κ intersecting the center C of the ellipse ε. A side view (a) and a 3D view (b) of the planar cone-section.

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.

A.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 4a, we derive R from d, θ and α as:

$$R = d\frac{\tan\theta}{\cos\alpha} \tag{2}$$

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:

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

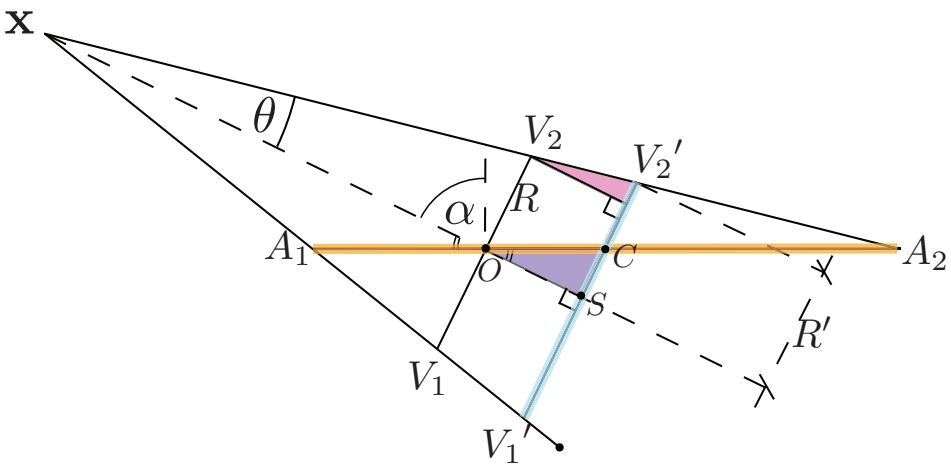


Figure 5: A detailed side-view of the intersection of ellipse ε and circle κ.

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

$$A = \frac{a_1+a_2}{2} \tag{5}$$

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 4b. The intersection line $\varepsilon \cap \kappa$ is perpendicular to A_1A_2 and intersects the center C of ε. Consequently, $\varepsilon \cap \kappa$ is collinear with the minor axis of ε. Figure 5 illustrates the side view of $\varepsilon \cap \kappa$. In Figure 6, we focus on the triangles $\triangle XV_2'V_2$ 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.

$$\sin\alpha = \frac{d'}{\|OC\|} \tag{6}$$

$$\tan\theta = \frac{dR}{d'} \tag{7}$$

$$R' = R + dR \tag{8}$$

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

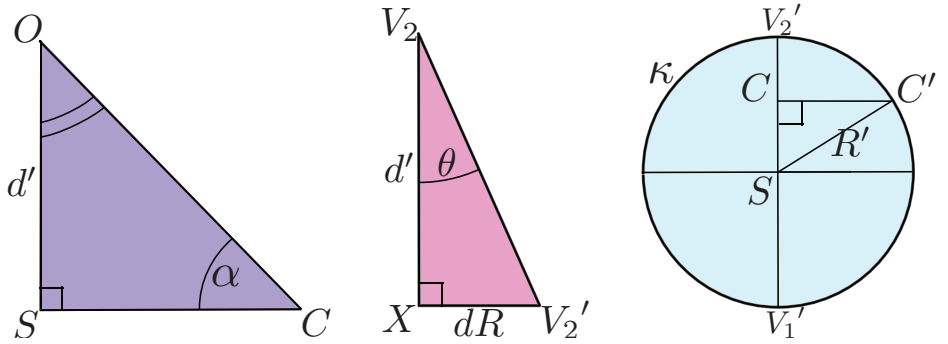


Figure 6: We introduce new literals for selected primitives: 6a the triangle $\triangle SCO$, 6b the triangle $\triangle XV_2'V_2$ and 6b the circle κ. These primitives are defined in Figures 4 and 5 by the same color encoding. Note, that $||CC'|| = B$ which is the minor axis of the ellipse ε.

A.3.3 Weighting Function

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

$$W_L(x,y) = 1-k \tag{10}$$

with k defined implicitly by:

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

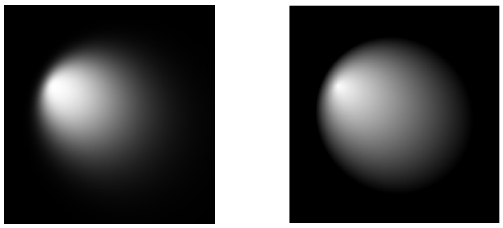
Equation 11 results in a quadratic equation with two real roots from which we take the maximum. A kernel with a linear fall-off from O towards the border of the ellipse is illustrated in Figure 7a. Additionally, we apply the Gaussian function to smooth the fall-off of the weights as illustrated in Figure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

A.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 **L'** and rotate the kernel so that its major axis is aligned with **L'**:

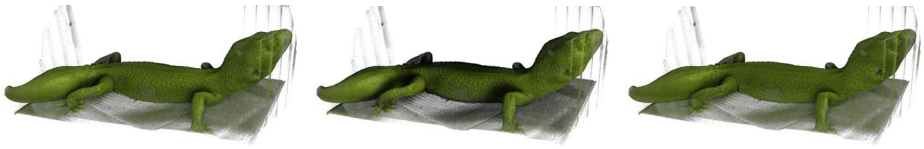
$$\frac{\vec{OC}}{\|OC\|} = \frac{\mathbf{L'}}{\|\mathbf{L'}\|} \tag{12}$$

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



(a) (b)

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).



(a) (b) (c)

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

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:

$$\lambda_i = \frac{1}{1 + G_\varepsilon * O_{i-1}} \tag{13}$$

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:

$$O_i = G_\varepsilon * O_{i-1} + \alpha_i \tag{14}$$

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 precomputation 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.

A.4 Implementation Details

Our new model was implemented as a plugin to VolumeShop [10] 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 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.

A.5 Results and Discussion

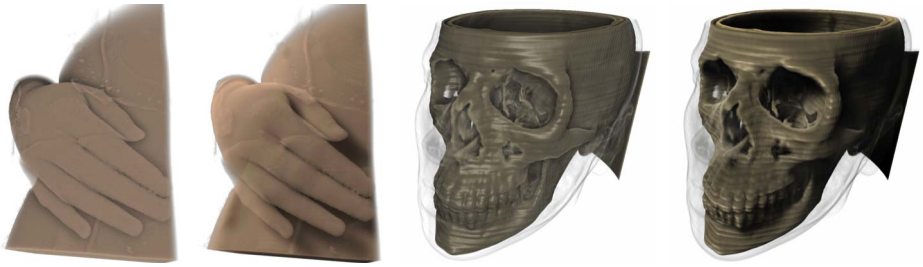
In this section, we provide case-studies and comparisons to other volume rendering approaches and analyze the performance of our new method.

A.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 9a, the image lacks depth, Figure 9b emphasizes the detailed structure of the skeleton through shadows. Similarly, Figure 10 shows cases where illumination leads to better perception of structures. In Figure 10a, the hand seems to directly contact the body. In reality, there is a small gap which is visible in Figure 10b. Similarly for Figures 10c and 10d: in Figure 10d, the eye sockets of the skull appear deeper than in Figure 10c. 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 10b and 10d. 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 11a and 11b use the multidirectional OS model enhanced by specular highlights, and Figures 11c and 11d 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

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



(a) (b) (c) (d)

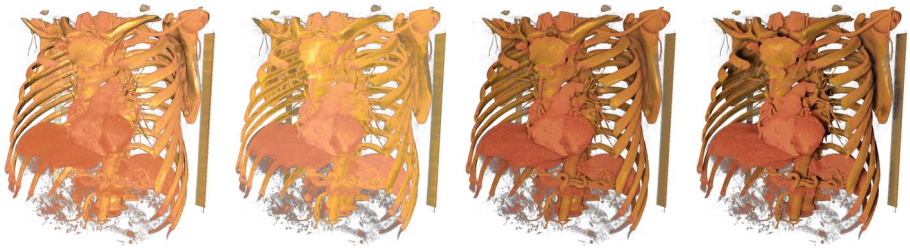
Figure 10: Visualizations of computer tomography data using the directional occlusion shading model with a headlamp illumination setup 10a and 10c using the illumination setup conventional to medical illustrations 10b and 10d.

(a) (b) (c) (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 11a and with the top-left lighting 11b. Visualizations 11c and 11d use the diffuse illumination model with the headlamp and the top-left light-source setup respectively.

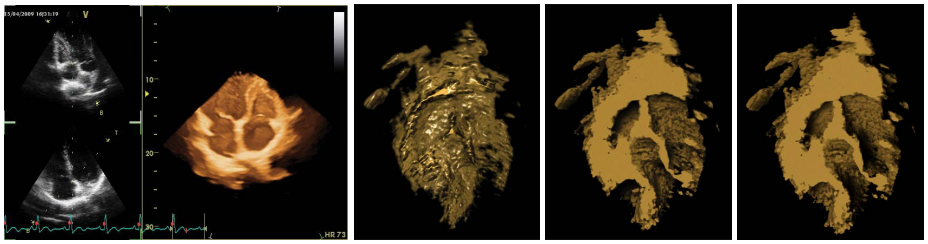
the hand, 39 participants (92.86%) favored the top-left illumination in Figure 1d, 2 participants (4.76%) preferred the raycasting in Figure 1b and 1 participant (2.38%) preferred the head-lamp illumination in Figure 1c. A majority of 41 (97.62%) participants also preferred the top-left illumination of the thorax in Figure 12d and only one participant (2.38%) selected the raycasted image in Figure 12a.

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



(a) (b) (c) (d)

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

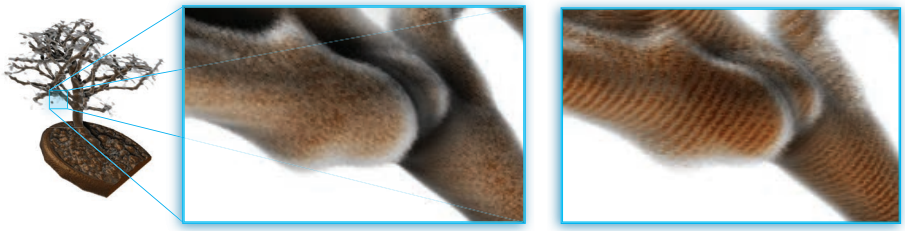


(a) (b) (c) (d)

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

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 13a. We used a transfer function which shows the heart in a similar fashion. Figure 13b 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 [54] or light volumes [127], 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. [70], 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



(a) (b)

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

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.

A.5.2 Performance Analysis

We tested the described method on a workstation equipped with an NVIDIA GeForce 295 GTX GPU with 1.7GB graphics 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 multidirectional 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.

A.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.

A.7 Acknowledgments

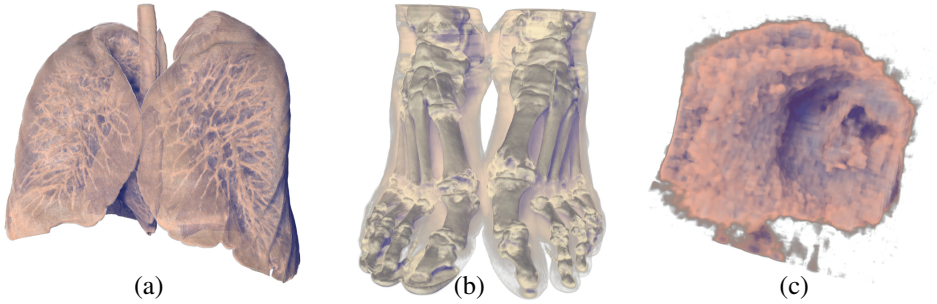
This work was carried out within the IllustraSound research project (# 193170), which is funded by the VERDIKT program of the Norwegian Research Council with support of the MedViz network in Bergen, Norway. The authors wish to thank the certified medical illustrator Kari Toverud for consulting, numerous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

PAPER

B

Chromatic Shadows for Improved Perception*

Veronika Šoltészová[1], Daniel Patel[2], Ivan Viola[1]



(a) (b) (c)

Figure 1: Non-photorealistic shadows expressed in blue tones: (a) lungs and (b) ankles from the visible female CT dataset, and (c) 3D cardiac ultrasound.

Abstract

SOFT shadows are effective depth and shape cues. However, traditional shadowing algorithms decrease the luminance in shadow areas. The features in shadow become dark and thus shadowing causes information hiding. For this reason, in shadowed areas, medical illustrators decrease the luminance less and compensate the lower luminance range by adding color, i.e., by introducing a chromatic component. This paper presents a novel technique which enables an interactive setup of an illustrative shadow representation for preventing overdarkening of important structures. We introduce a scalar attribute for every voxel denoted as *shadowiness* and propose a shadow transfer function that maps the shadowiness to a color and a blend factor. Typically, the blend factor increases linearly with the shadowiness. We then let the original object color blend with the shadow color according to the blend factor. We suggest a specific shadow transfer function, designed together with a medical illustrator which shifts the shadow color towards blue. This shadow transfer function is quantitatively evaluated with respect to relative depth and surface perception.

* This article was published in *Proceedings of the Symposium on Non-Photorealistic Rendering and Animation (NPAR 2011)*, 105–115, 2011 and presented at named symposium in Vancouver, Canada by Veronika Šoltészová. A typo in definition of shadowiness $\mathscr{S}$ has been corrected in the first paragraph of Section B.3.

[1] University of Bergen, Norway

[2] Christian Michelsen Research, Norway

B.1 Introduction

Shadows are important monocular and also binocular [123] depth cues. Based on the shadows, the observer can estimate the relative depth between the foreground object and the background, and estimate the spatial ordering of occluding structures. When no shadows are being cast, objects in a rendered scene often appear floating in the space. Thus the realism of a scene increases significantly when shadows are cast.

During image synthesis, shadows can be computed accurately with physically-based global illumination methods. Such rendering algorithms come with high computational requirements and only simple scenes can be synthesized at interactive frame-rates. To achieve interactive performance for volumes or complex meshes, shadowing approximation algorithms can be utilized. One category is approximations of global illumination light propagation and the other category is dedicated shadow mapping algorithms that calculate shadow geometry for each object in the scene.

High computational cost is not the only limitation of shadows. During data visualization, the user should be provided with the best possible visual information. Here shadows can result in unwanted information hiding as they darken the scene elements in the shadow. A visualization user might thus receive intuitive depth cues, but due to overdarkening caused by the shadows he or she might miss important features. Therefore careful consideration should be taken when introducing shadows in 3D data visualization scenarios.

The craft of illustration has been a great source of inspiration for visualizing data. In the case of shadows, illustrators have tackled the problem of overdarkening. When a shadow is cast, the area of an object in shadow is a darker color than the area outside the shadow. Thus the logical way to modify the color of an object, based on the degree of shadow, is to modify the luminance of the color. However, the illustrators approach to tackle overdarkening is different from physically-based approaches that assume white light. They change the color less in luminance but add a chromatic component so that the area in shadow is still visible but the added coloring conveys that the area is in shadow. Often, a blue tone is used for the shadows, however, also a complementary color to the object's material can be used. Blue color has been used to express shadowing in several artworks such as still lifes by Cézanne [16, 17], paintings from the *haystack*- and *Rouen cathedral* series by Monet [101, 102], the self-portrait by Matisse [91] and landscapes by Coleman [24]. It is interesting to compare the two versions of da Vinci's Virgin of the Rocks shown in Figure 2. Compared to the painting [30] in Figure 2a, da Vinci used more blue pigments in the later painting [31] in Figure 2b. The blue tones do not only enhance the aerial perspective but allows to see more details in the shadowed areas. Instead of using black as in Figure 2a, different blue shades are used in the crevices of the dress of the Virgin and also in the shadow of the rocks in the distance.

Blue shadows imitate what is observable on a cloud-free sunny day, where objects in shadow appear to be blue-toned. This is because the blue sky dome acts as an indirect light source which is not perceivable on the object directly exposed to the sun light, but becomes noticeable for an object in the shadow. Especially on white reflective materials, such as snow, the blue tone in shadow is clearly noticeable. Due to the blue sky dome, humans might be used to observe objects in the shadow as having a subtle blue tone and it is possible that illustrators use this effect to overcome shadow overdarkening.

(a) (b)

Figure 2: Leonardo da Vinci: Virgin of the Rocks. (a) The version from 1483–1486 located in Louvre, Paris. (b) The version from 1495–1508 located in the National Gallery, London.

Inspired by the shadowing technique from the illustration craft, our goal is to provide a non-photorealistic shadowing approach for 3D data visualization that handles overdarkening. For each sample we therefore extract the amount of shadow, i.e., *shadowiness* and map it via the *shadow transfer function* to a color and a blend factor. Based on the amount of the *shadowiness*, the user can flexibly control the visual appearance of the shadow. In Figure 1, the technique is demonstrated on medical datasets. A flexible approach to interactively control the optical properties of the shadow is the first contribution of the paper.

During an informal feedback, our chromatic shadowing method was well received by illustrators. We were also interested in how the blue illustrative shadows performed in terms of shape, contrast and depth perception compared to standard black shadowing based only on luminance shift. Therefore we evaluated how subjects performed on perception tasks in both scenarios. Based on the conclusions we draw from the evaluation, we suggest an optimal choice for the shadow color in terms of both a shift in chromaticity and luminance. This is the second contribution of the paper.

B.2 Related Work

The human visual system is fine-tuned for inferring shape by observing how light interacts with physical objects. Artists were early in recreating, in pictorial form, some of the drawing cues that the brain infers shape from [49]. With the advent of optics and computer graphics, automated algorithms for creating shape-cues is being explored. These algorithms have both objective and subjective aspects. Examples of objective aspects are which optical phenomena they recreate and how computationally expensive they are to calculate. Subjective aspects can be which type of cues they give to the brain (distance, curvature, motion) and how well the brain can make use of these cues. There exists optically complex phenomena which are computationally expensive to calculate, such as polarization [161]. They contribute towards photorealism, but give few shape cues, if at all. On the other hand, using gradients for local shading on surfaces [117], and volumes [80] produce strong shape-cues which in addition are fast and simple to calculate. They are therefore widely used in interactive renderings. Shadows are also strong cues, but due to their global nature, are more time consuming to calculate. Since the brain infers shape well from less accurate shadows [156, 168], many shadow rendering techniques can be accelerated by reducing their accuracy. In the following we give an overview of raytracing methods and approximative but fast shadow algorithms.

Yagel et al. [166] employed recursive ray tracing for correct shadows. Behrens and Ratering [5] added shadows to texture-based volume rendering by using an additional shadow volume. Rezk-Salama [125] proposed a semi-interactive approach for GPU-based Monte Carlo volume raytracing. Faster shadow calculations were achieved by limiting the occlusion calculation to local neighborhoods; Ambient occlusion, where 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, was described by Zhukov et al. [169]. Dachsbacher et al. [32] treat visibility implicitly while propagating obscurance as an additional quantity. These approaches are view-independent, therefore, for fixed geometry, occlusion information only has to be computed once and can then be applied efficiently during rendering. Hernell et al. [56, 57] present interactive ambient occlusion by coarse neighborhood sampling. This approach requires ambience data for each voxel to be recomputed when changing the transfer function. Knecht [69] and Méndez-Feliu and Sbert [96] provide overviews of rendering techniques based on ambient occlusion.

Some shadow approaches do not require extensive precomputation and therefore allow fully interactive transfer function modification. The model presented by Kniss et al. [70, 71] captures volumetric light attenuation effects using half-angle slicing. Schott et al. [134] improves the speed of the half-angle method by aligning the light source direction with the viewing direction. Speed-up is achieved by calculating the light attenuation in lock-step with the standard slice-based front-to-back traversal of the volume for compositing transparencies. The result is that the radiance for each voxel is now based on a conical neighborhood pointing towards the viewer, extending to the end of the volume. These methods therefore create view dependent shadows. Šoltészová et al. [139] generalized the latter work for user-defined light source directions.

We recreate the effect of colored shadows seen in the works of artists and medical illustrators. The concept is specified by introducing a user defined chromatic shadow transfer function that gives control of color gradation in shadows. We achieve realtime

shadows on volumetric data by building on the work by Šoltészová et al. [139]. The use of artistic-coloring techniques in computer graphics has been performed earlier. By using a colored shadow, we avoid losing detail in shadow areas of high darkness. A similar motivation led Amy Gooch et al. [50] to recreate an illustrative rendering style for geometric objects by mapping the local Phong grayscale shading intensities to a color scale. As an extension, Bruce Gooch et al. [52] added silhouette curves, improved shading, and simple shadow cast from geometry to a flat plane for conveying shape better. Inspired by colored shadows found in illustrations, they provide an example rendering with a colored shadow, without going into more detail. Ruiz et al. [130] present cool-to-warm mapping on ambient occlusion. Their method requires precomputation for the scene. None of these papers quantify the effect with user studies as we do.

By user studies, our paper compares the perceptual effect of shading techniques. The effect of shading on shape was performed by Ramachandran [124]. He found that when extracting shape information, the brain assumes that there is only one light source illuminating an object. In our model, we operate with one light source. O'Shea et al. [109] use local Lambertian shading and investigate how good a shape-cue this is. They find that shape is perceived best when the light direction is above the viewpoint at an angle, this is a setup which we use in our model. In one study [159] it was shown that global illumination has perceptual benefits over local illumination. Winnemöller et al. [163] evaluate the shape-conveying effect that different non-photorealistic techniques have. Specifically, they measure the effect of shading, texture, silhouette and motion on the perception of shape. Besides all, they showed that when multiple shape cues are present in moving objects, the cues may conflict and impede each other. For simulating global illumination, Krivanek et al. [76] study how many point light sources are needed to get an acceptable visual correspondence to a ground truth rendering.

Ware and Franck [158] performed user studies to find the effect of 3D stereo viewing and motion on depth perception. Pfautz [115] discusses stereo and perspective depth cues. Grossman and Balakrishnan [53] compared depth perception on a volumetric display. Successful use of textures for conveying shape was investigated by Bair et al. [3].

B.3 Transferring Shadowiness to Chromaticity

In computer graphics, fast shadow-approximating techniques use multiplication of the color: r,g,b $\in [0,1]$ by the degree of their shadowiness $\mathscr{S} \in [0,1]$. $\mathscr{S} = 1$ means total obscureness where colors become pure black and $\mathscr{S} = 0$ signifies no shadow at all; colors remains unmodified. As a consequence, shadow areas become darker and thus the depth cues are enhanced. Unfortunately, the darkening is controlled only by the scene and light setup. It cannot be assured that the shadowed areas do not become completely black and mask out all visual information. When no shadowing is applied to the scene, visual information in such areas such as local surface detail will remain, however the depth cues from shadows are lost. In volume visualization, both aspects are important: the amount of information being communicated from the image is crucial and also depth cues for better 3D-scene understanding.

We observed that illustrators tend to avoid expressing shadows by mixing the object color with a pure black color. Our advising medical illustrator indicated that blue color is

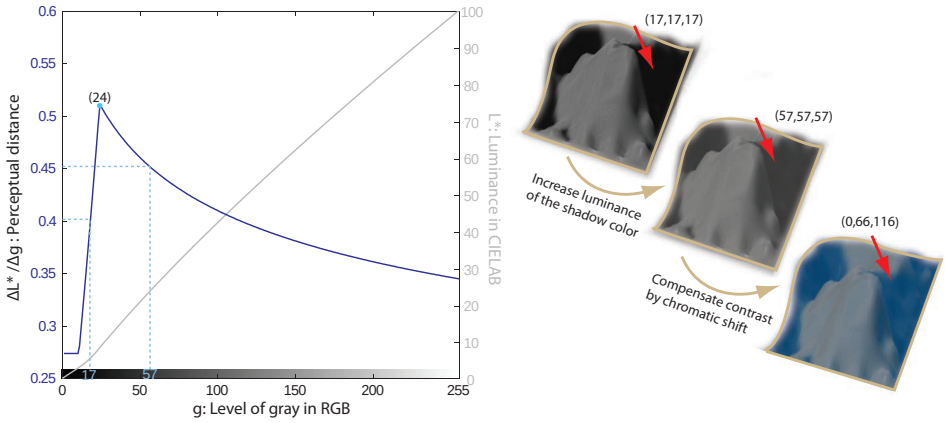


Figure 3: The non-linearity of the RGB grayscale: the grayscale colors converted to luminance (L*) in the perceptually uniform CIELAB color space with the scale on the right and $\frac{\Delta L^*}{\Delta g}$, the perceptual distance between neighboring grayscale colors in the RGB space with the scale on the left. We compare a shadow with a very dark gray color CIELAB = (5,0,0) / RGB = (17,17,17), a brighter shadow color (24,0,0)/(57,57,57), and a chromatic shadow (24,-16,-39)/(0,66,116) of the same luminance level as the brighter shadow. The arrow point to the region where features are visible only in the chromatic shadow.

an optimal choice. In other words, she partially uses luminance contrast, i.e., darkening, and partially chromatic contrast, i.e., mixing in a blue tone.

This can be explained by the fact that the human visual system does not distinguish between individual colors in the lowest luminance range. In the perceptually uniform CIELAB color space, the Euclidean distance between color vectors corresponds to their perceptual distance. We converted gray values from RGB color space to CIELAB. As the grayscale only varies in the L*-dimension (luminance), it is sufficient to regard the first derivative of L* to show the perceptual distance between neighboring gray colors. In Figure 3, we plot the RGB grays in the interval [0,255] converted to L* and the derivative $\frac{\Delta L^*}{\Delta g}$. We observe that the derivative peaks in the low-middle range of the grayscale axis in the RGB space which corresponds to the highest perceptual distance. Therefore it is logical to use the more effective range of gray in terms of perceptual distance instead of the pure black and use the chromatic shift to compensate for the decreased luminance contrast in the artwork. Shadows encoded solely in the luminance channel do not have the same contrast as shadows encoded with variation in luminance and chromaticity.

Through an adaptation of the fast shadowing methods for mimicking artistic illustration techniques, we combine the best of both scenarios: depth cues from shadows and detail extraction when no shadowing is used. We propose a method which gives the user full control over the shadow representation in terms of steering its color and opacity and allows for mimicking the illustrative shadowing technique. A suitable choice of the shadow color allows to communicate more local features in the shadowed 3D scenes.

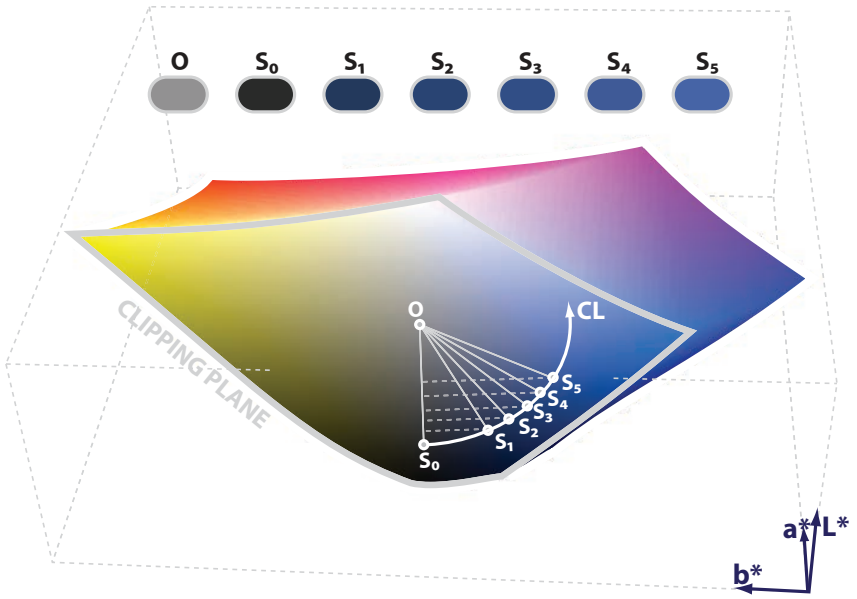


Figure 4: Palette of shadow colors $S_0..S_5$ in the CIELAB space: O is the object color and shadow colors $S_0..S_5$ are chosen along the planar arc CL. CL denotes the chromatic-luminance shift and belongs to the circle in the CIELAB color space which is centered in O and has a radius $||OS_0||$. We used the plane define by CL to clip a part of the CIELAB volume to reveal the color plane of our interest.

Volume-based chromatic shadows

As we focus on visualization of volumetric data, our starting point is a volumetric renderer which is capable of generating a soft-shadowing effect. Howevever, any shadowing algorithm that outputs a scalar shadowiness value can be used.

With the shadow transfer function f, we map the shadowiness $\mathscr{S}$ to a color C^s_{RGB} and a blend factor C^s_A: $f(\mathscr{S}) \rightarrow \mathscr{C}^s_{RGBA}$. The original voxel color $\mathscr{C}^o_{RGB}$ is blended with the shadow color $\mathscr{C}^s_{RGB}$ according to the blend factor C_A to obtain the final voxel color $\mathscr{C}_{RGB}$:

$$\mathscr{C}_{RGB} = (1 - \mathscr{C}^s_A)\mathscr{C}^o_{RGB} + \mathscr{C}^s_A \mathscr{C}^s_{RGB} \tag{1}$$

The function f is user-specified and it allows with a suitable choice of color to divide the *contrast energy* of $\mathscr{S}$ into luminance- and chromatic contrast. In Figure 3, we show the same part of an artificially created scene with three shadow presets. The dark makes it difficult to perceive the surface shape in the dark crevices. The perceived color distance of the original surface color, i.e., CIELAB = (60,0,0), to the shadow color is 55. Then we increased the luminance of the shadow color, but the perceived color distance to the original surface color decreased to 36 and the image lost its contrast. Then we compensate for the loss of contrast by chromatic shift of the shadow color into the tones of blue. We chose a color which has the same perceived color distance to the object color as the color of the dark gray shadow.

In Figure 4, we present another example of the shadow palette on a visualization of the CIELAB volume. Shadow color $S_0..S_5$ are equidistant from the object color O. With

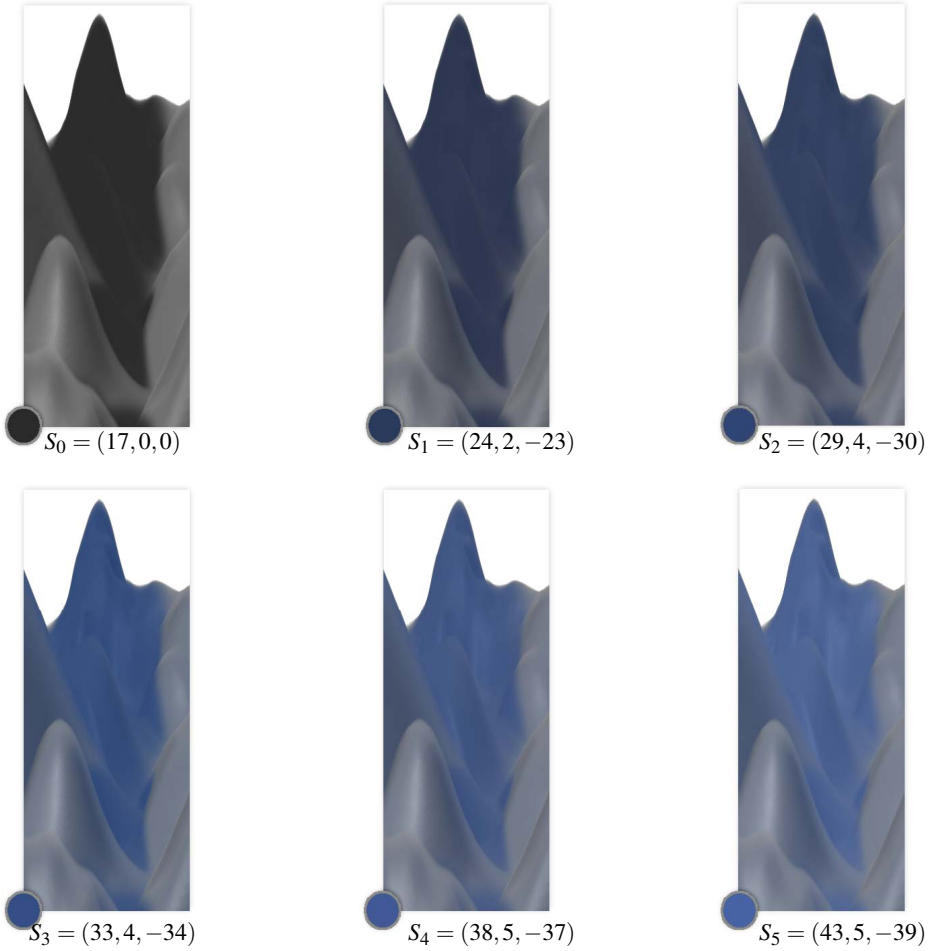


Figure 5: A volumetric dataset generated from several sine waves. Using a step transfer function, we visualize one surface in the color O. We show how the shadow colors $S_0..S_5$ influence the surface perception in the shadow. O and $S_0..S_5$ are from the palette in Figure 4 and are defined in CIELAB.

growing i, S_i contains less luminance contrast (ΔL^*) and more chromatic contrast (Δa,Δb) to the object color. An application of individual shadow-colors from this palette is showed in Figure 5.

Image-based chromatic shadows

In theory, chromatic shadows can be generated in a post-processing step from two renderings: one with and one without shadows. In this step, the renderings are compared pixel-by-pixel. From the difference between luminance of individual pixels, one can find the shadowiness $\mathscr{S}$, apply $f(\mathscr{S}) \rightarrow \mathscr{C}^S_{RGBA}$, and blend the shadow color with the original

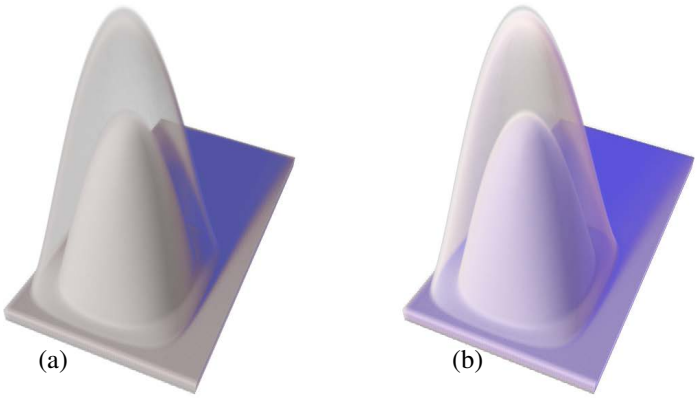


Figure 6: Comparison of chromatic shadows on a phantom volume generated (a) image-based as post-processing and (b) volume-based during the volume rendering pass. We show that the image-based method does not deal with the transparency correctly. The shadow color is applied as post-processing (a) and cannot appear behind a semi-transparent surface as in (b).

pixel color in the rendering without shadows using Equation 1.

However, this image-based method has an important shortcoming: it does not handle transparencies correctly. This is displayed in Figure 6a. We created an artificial dataset where a semi-transparent thin layer encapsulates an opaque object on a flat board. It is obvious that the shadowing on the surface does not appear to be drawn behind the semi-transparent layer as it is in Figure 6b, which was rendered using the volume-based approach. In addition, the volume-based approach also captures a subtle shadow on the back-face of the semi-transparent layer cast by the opaque object located inside. It was not our goal to show both figures with a corresponding coloring of the shadow, but to highlight an important disadvantage of the image-based method. In volume visualization, we need to handle transparencies correctly which justifies the choice of the volume-based shadowing method.

B.4 Results

Our method is beneficial for volume rendering of data from several scientific domains. The seismic dataset in Figure 7 contains many layers. Rendered with black shadows, the space between layers becomes overdarkened and illegible as shown in Figure 7a. The surfaces of the layers are better visible with our method as shown in Figure 7b. In Figure 8, we show two pairs of medical visualizations comparing black shadows to our method. Other example applications of our method for medical datasets are shown in Figure 1. The lungs and feet are extracted from the visible female CT dataset (Figure 1a and 1b). Figure 1c shows a 3D ultrasound volume of a human heart.

The non-photorealistic shadowing technique we described in this paper belongs to the group of low-frequency lighting approximations and as suggested in previous works, low-frequency lighting is suitable for visualization of data with low signal-to-noise ratio

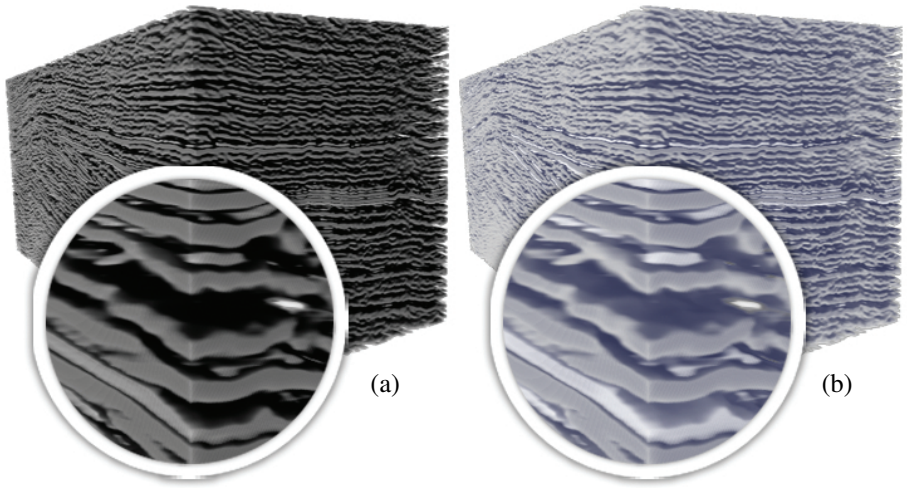


Figure 7: A seismic dataset rendered with (a) black and (b) blue(chromatic) shadow color.

such as medical ultrasound or MRI [57, 127, 139] and seismic datasets [112]. Features in shadows are illegible, especially in the medical domain. Figures 7 and 8 demonstrate that our non-photorealistic shadowing technique brings a substantial improvement over the previous shadowing techniques in terms of feature visibility.

We implemented a proof of concept of the described non-photorealistic shadowing technique. We used the volume rendering framework VolumeShop [10] which is written in C++ and uses OpenGL/GLSL. We chose the multidirectional occlusion shading model [139] since it is capable of generating a high-quality soft shadowing effect at interactive frame rates and therefore suited for interactive generation of illustrations. In addition, this model allows us to place the light source anywhere on the hemisphere behind the viewer. The chosen model supports only volumes but the concept of chromatic shadows should fit well for meshes also.

The multidirectional occlusion shading model requires a slice-based volume renderer. It computes shadows using incremental convolution: for each slice the volume samples are multiplied with a 2D occlusion buffer generated at the stage when the previous slice was rendered and a new occlusion buffer is produced for the next slice. Instead of multiplying the color of samples $P(x,y)$ by the values from the occlusion buffer $O(x,y)$, we interpret them as shadowiness $\mathscr{S}$ and use to specify the shadow color.

The mapping function of the shadowiness to color and opacity is user-defined as a 1D transfer function. Currently, users specify the colors in RGB. Automatic shadow color specification based on the voxel color can be seen as a possible extension of this method. Colors defined in CIELAB which were discussed in this paper were converted separately. The LAB $\leftrightarrow$ RGB conversion was done using the standard illuminant D50 with the following XYZ reference values $\mathrm{XYZ}_{ref} = (96.4212, 100.0, 82.5188)$. This is the same as used by Adobe Photoshop in the image editing mode [89].

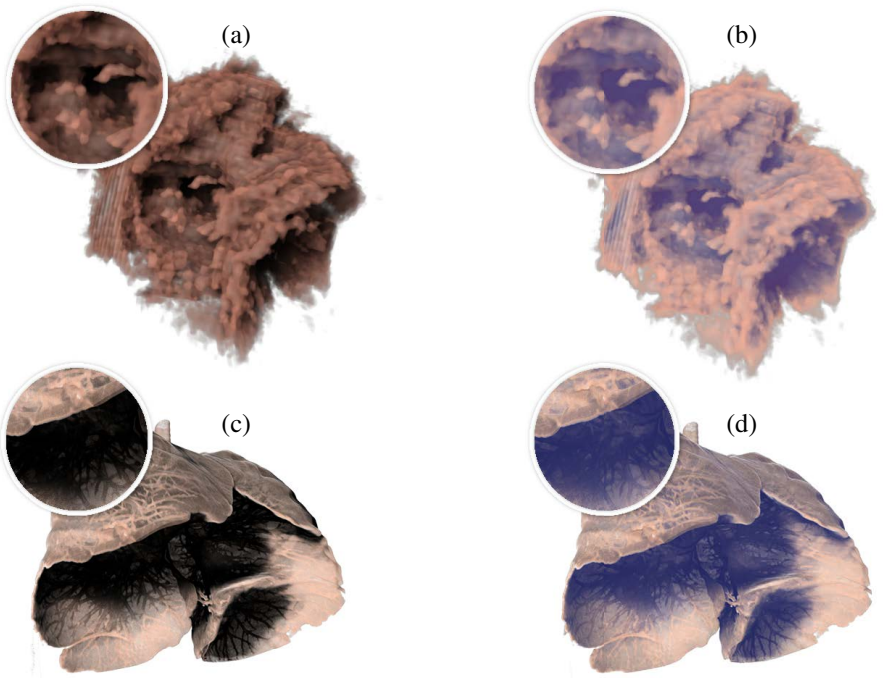


Figure 8: Visualization of medical data: cardiac ultrasound using (a) a black and (b) a blue shadow color, lungs from a CT dataset rendered with a (c) black and a (d) blue(chromatic) shadow color.

These reference values were used to generate a CIELAB gamut volume shown in Figure 4. Compared to the previous multidirectional occlusion shading model, our method requires a shadow color definition and that an additional color-mixing operation is performed for every sample during volume rendering. This operation did not degrade the performance of the application and we achieved the same frame rates as the referenced work.

Initially, during informal meetings with an illustrator, when showing visualizations with graylevel shadows, the illustrator commented that the images “look dirty”. With standard shadow rendering, we had no control over the appearance of the shadow. We followed the suggestions of the illustrator until we obtained satisfying visualizations which utilized shadow colors of darker blue tones such as 1c and 8b. These results were driven by a subjective quality assessment of the illustrator. Later we will give a statement about the quantitative perceptual benefit of our method.

B.5 Validation

Our hypothesis was that an illustration-inspired shadow specification enhances the perception of shape in shadow and allows to better distinguish features as compared to traditional shadowing techniques. Potentially, the modification of shadow might also affect

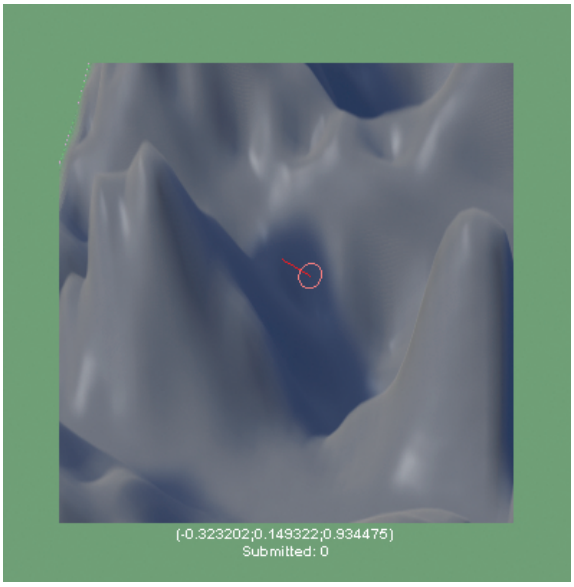


Figure 9: An example screen capture from the surface-perception task. The gauge figure is drawn in red.

depth perception. To quantitatively test our statements, we studied shape and depth perception in three experiments. This allowed us to observe trends and dependencies in the human perception of the scenes as a function of the darkness of the blue used for the shadow. First, we measured the accuracy of perceived surface normals in shadowed regions to prove our hypothesis that better visibility in the shadow will trigger more accurate normal estimation. Second, we conducted a contrast test, where we measured and compared the just noticeable variation of colors in images generated with different shadow parameters. Finally, we studied relative depth estimation between points in the visualization with luminance-only and chromatic shadow to evaluate the depth cues from the image.

B.5.1 Surface Perception

To measure the accuracy of surface perception in images, we designed a gauge-figure task. A gauge figure consists of a disk centered around a line and the user-task is to adjust the orientation of the figure so that the disk is perceived to be tangent and the line to be orthogonal to the surface. This task has been employed in previous work [72, 109] to quantitatively assess surface perception from images. Figure 9 shows an example screen capture from our experiment.

Setup

Each subject of the experiment was presented to a surface rendering which we artificially generated from sine-waves. The wave composition was complex so the underlying func-

tion could not be deduced from the image. In addition, we created a synthetic aperture around the regions of interest to guarantee that the subjects did not try to extrapolate the surface based on the whole scene. The aperture, .i.e., the window, had a rectangular shape and not circular to avoid that the subjects mistook the aperture for the silhouette of the object which would affect the results. We chose an artificially generated dataset for two reasons. First, all prior knowledge of the surface can be excluded. Second, we are able to calculate the ground-truth normals by analytic derivation of the function which we used to generate the scene instead of finite differences.

We used the guidelines for drawing the gauge figure from the previous work [109]: the gauge figure was rendered as a red line and a red disk using a perspective projection and the depth buffer was cleared before rendering the gauge figure. Otherwise, the figure could be partially occluded by the surface in locally convex areas. Both the line- and the circle-linewidth were one-pixel wide to minimize occlusion. The locations of the gauge figure were chosen randomly until a location in shadow was found. The subjects had to adjust the orientation of the figure by rotating it around its fixed point using a left mouse button. Subjects submitted the solved test case with the enter button and a new random location was chosen. The orientation of the gauge figure was always reset to default when it appeared at a new location. In addition, we allowed the user to scale the gauge figure using the right mouse button and to toggle different modi: line only (no disk) and fixed point only, i.e., one red pixel on the surface showing the fixed point. The diverse toggle modi allowed for less occlusion of the surface by the gauge figure. We did not allow positions of gauges at surface contours, i.e., locations where the angle between the surface normal and the viewing vector is almost 90° to avoid ambiguity cases where the location belongs to the region closer to the viewer or not.

Previous work [109] shows that light direction influences the accuracy . Therefore, the illumination was fixed to be from top-left in all test cases.

For each surface, we tested the shadow colors from the palette shown in Figure 4. In total, we utilized images of the same scene but from three different view-points so that each surface had approximately the same level of difficulty. For each view setup, we generated a series of images with varying shadow color. Figure 5 includes the examples of one series with constant view-point and with varying shadow color but constant shadow-to-object contrast. We restricted our experiment to the first five colors in the palette $S_0..S_4$.

Subjects and procedure

We performed the experiment on eight male subjects. Each subject was familiar with 3D applications and computer graphics and had no prior knowledge of the hypothesis. All subjects understood the notion of a *surface normal*. If they needed an optical correction, e.g., glasses or lenses, they wore it during the experiment. None of the subjects were color blind.

A session for one subject was divided into five subsessions, for each of the shadow colors $S_0..S_4$ in random order. All subjects were confronted with all shadow color being studied. In this way, we reduced the error induced by the overall ability to solve this task of individual subjects. We always chose a scene from a different view-point for subsequent sessions in order to avoid the subject remembering the scene from the previous subsession. We involved at least a half-day break after the third session so that we did not show the

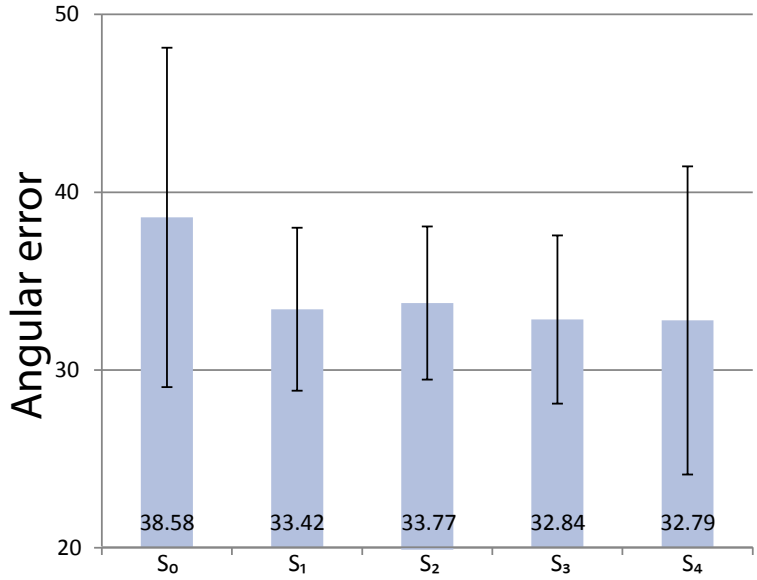


Figure 10: Angular error of an average subject for shadow colors $S_0..S_4$. The error bars represent the standard deviation.

scene from the same view-point in subsequent sessions and also to avoid tiredness of the subjects. All sessions were performed on the same workstation on a monitor calibrated by a *Spyder3* device and under constant lighting conditions in the room.

Each subject solved at least 20-25 test cases, i.e., estimated surface normals, per image. In total, we obtained 175-211 solved test-cases, i.e., samples, per shadow color and 1005 samples in total.

Results

For each test case, we calculated the angle between the estimated and ground-truth normal, i.e., angular error in degrees. We performed the angle calculation in the coordinate system of the viewer. For this reason, we transformed the ground-truth normal, originally in the coordinate system of the volume, using the inverse-transpose viewing matrix. The comparison yielded an average error for each shadow color being tested.

Based on the scores for individual shadow color of each subject, i.e., average angular error, we calculated the score of an average user for each shadow color shown in Figure 10. The results of our small-scale study of shape perception reveal a trend of improved shape perception towards brighter shadow colors under chromatic shift. The most significant improvement is between S_0 and S_1.

B.5.2 Contrast Perception

According to the previous work [109], some of the errors we measured in the experiment with the gauge figure can be attributed to the task itself. Therefore, to support our

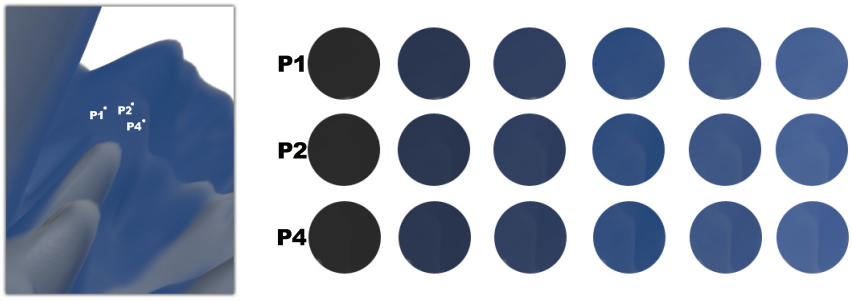


Figure 11: A screenshot of the scene with shadows S_3 used for the contrast experiment. For three selected points, P_1, P_2, P_4, we show an aperture with 45-pixels radius for the series of shadow colors, $S_0..S_5$ depicted from left to right, that we tested in the experiment.

statement regarding the perceptual improvement using the non-photorealistic shadows, we conducted an additional experiment. For selected shadow colors, we compared sizes of smallest possible circular areas centered around a selected point where a subject is able to perceive a just-noticeable difference.

Setup

We utilized the same dataset as for the gauge-figure task as shown in Figure 11. We investigated six different shadow colors $S_0..S_5$. We revealed only a part of the scene through a circular aperture with the radius of 3 pixels centered around a selected point in the image. In each test case, the subject enlarged the radius of the aperture until he or she could see a just-noticeable color variation through the aperture.

We chose seven different positions in the image for which we tested the aperture. The points were chosen from locations in shadow and such that an aperture centered at these points with radii 50 pixels reveal only such pixels which are attributed to the scene and not to the background. Figure 11 shows three out of seven position which we studied. For each point, the figure also shows a series of apertures for the whole palette of shadow colors which we investigated in this experiment. After one test case was solved, a new point was randomly selected from the set of pre-selected points in conjunction with a randomly set shadow color.

Subjects and procedure

This experiment was performed on eight subjects. Six of them had computer-graphics background and two had not. Seven subjects were male and one female. All subjects who required an optical correction wore it during the experiment. The testing was performed at the same calibrated monitor under constant lighting conditions. We explained the task first. Then subjects were solving the test cases one after another. Between individual test cases, we showed a gray screen for two seconds to neutralize the after image of the visual stimuli.

In total, we gathered 500 solved test cases from all subjects.

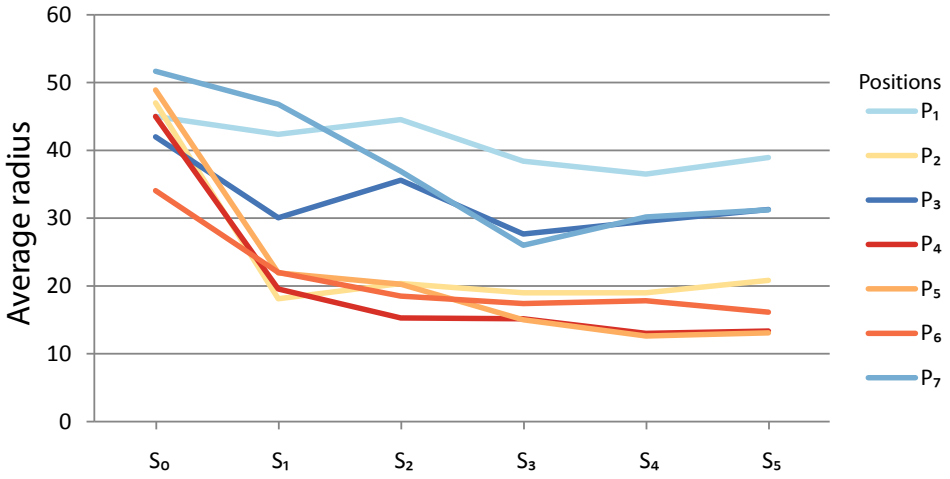


Figure 12: Average aperture radii for shadow colors $S_0..S_5$ at positions $P_1..P_7$.

Results

First, we compared average radii for each shadow setup for individual positions of the aperture $P_1..P_7$ as shown in Figure 12. We observed four strong (P_2, P_4, P_5, P_6), two weak (P3,P7) and one very weak (P1) decreasing trend. In Figure 11, P1, P2, P4 refer to aperture positions P_1 with a very low decreasing tendency, P_2 and P_4 which decreases faster in the diagram shown in Figure 12. The plotted series of apertures in Figure 11 clarifies the situation – the features around the aperture position P_1 are very weak compared to P_2 and P_4. Logically, the benefit in feature perception from our technique is smaller in areas where only subtle features are present.

We have therefore additionally separated aperture positions into two groups: G_1=(P_2, P_4, P_5, P_6) with strong features in the neighborhood and G_2=(P_1, P_3, P_7) with weak features. Figure 13 compares average radii for individual groups and for all samples. The samples within G_1 also have a smaller standard deviation which indicates a higher aggregation of the estimations while the estimations within G_2 (weak features) seem to be more dispersed. In favor of brighter shadows, some of the subjects commented that they were confused whether they were seeing a weak feature or an illusion. The large deviation can be due to that the subjects solved these ambiguous cases differently. S_0 has lower deviation within G_2 because most subjects agreed at the same large radius. As for the strong features, we obtained comparably lower standard deviations and a large difference in favor for brighter shadows with a chromatic shift.

B.5.3 Depth Perception

Shadows in computer graphics are used to enhance depth cues in images. As we modify the traditional representation of the shadow with the goal of improving the shape perception in shadowed areas, it is a question as to what extent the shadow modifies the original

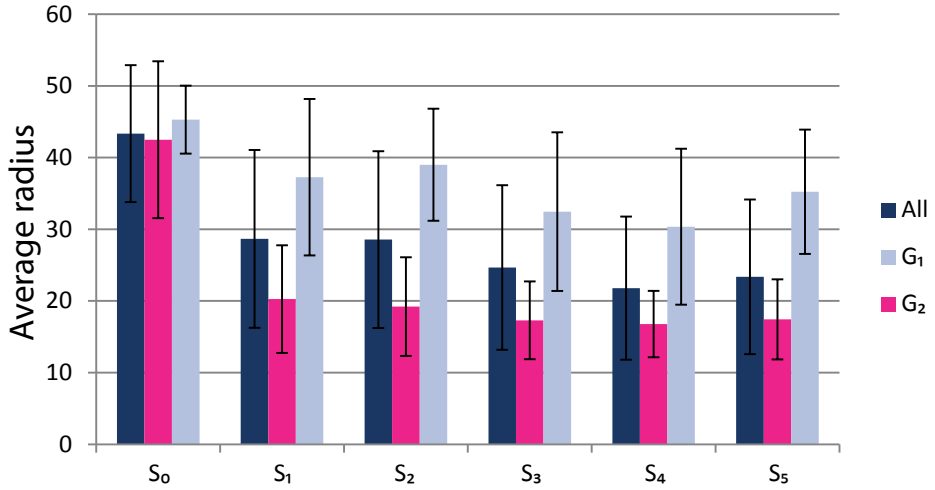


Figure 13: Average aperture radii for individual shadow colors and their standard deviations (error bars) plotted for all positions, for a position with strong features in the neighborhood G_1 and weak features G_2.

depth cues. Therefore, we conducted an experiment, where we compared the accuracy of relative depth estimation for three different lighting effects: the black shadow approach with decreased luminance in the shadow, illustrator specified shadow and no shadow (only local Phong illumination). The subjects compared the depths of three points shown on the image.

Setup

As a stimuli, we chose medical datasets acquired from 3D cardiac ultrasound. Tissue scanned with ultrasound has a low signal-to-noise ratio and therefore, its surfaces cannot be easily identified without training. The group of tested subjects included no clinical experts in ultrasound, therefore we could exclude estimations based on prior knowledge.

We used in total four different scenes of ultrasound heart and applied three different lighting effects to each of them. Light direction was constant for all lighting effects.

Procedure

This experiment was distributed to a wide number of people and 63 subjects participated in this experiment. The tested group included 50 experts in computer graphics. The non-experts included an artist, medical students, two pilots, and others. 55 subjects were male and eight female. One subject was red-green color-blind.

Each subject was exposed to four ultrasound visualizations rendered with either Phong shading without shadows, with an original soft shadowing effect using only darkening in shadow, or with illustrative chromatic shadows. Three points appeared over the visualization: a red, a blue and a yellow point as seen in the screen capture in Figure 14. The yellow point was crossed in order to add an additional cue for yellow-blue color-blind

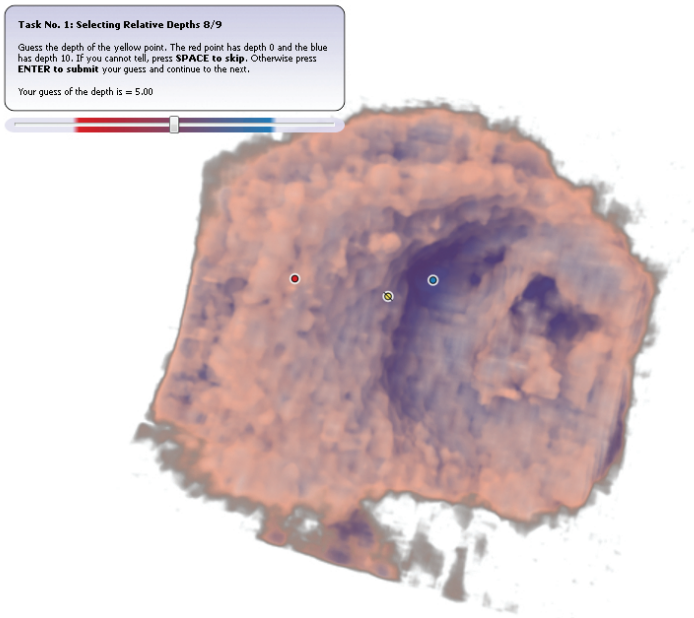


Figure 14: An example from the testing of depth perception. Subjects were estimating the relative depth of the yellow crossed point with respect to the red and blue slider and submitted their answers using the slider.

subjects. The small graphics overlay instructed the subjects that the red point is located at depth = 0 and the blue point at depth = 10 and to adjust the slider to the depth of the yellow point defined by the red and blue points.

Each respondent downloaded the application and performed the experiment at his or her workstation. The application gave clear instructions for the experiment and after the experiment finished, it produced a logfile which the respondend returned via email. For each of four possible scenes with a random selection of lighting effect, subjects solved two such tasks. In total, we received 506 solved test cases.

Results

We regarded the absolute difference between the estimated and ground-truth relative depth as the relative error. The ground truth was calculated with values from the depth buffer containing depths of the first non-transparent sample. The depth range we indicated in the instructions, i.e., $[0, 10]$ was scaled to the interval $[0, 1]$. Errors greater than 1 signify that subjects perceived the points in wrong depth order.

Our analysis is summarized in Figure 15. Even though the depth cues from blue shadows are slightly degraded as compared to the black shadows, they are still significantly superior than the local Phong illumination.

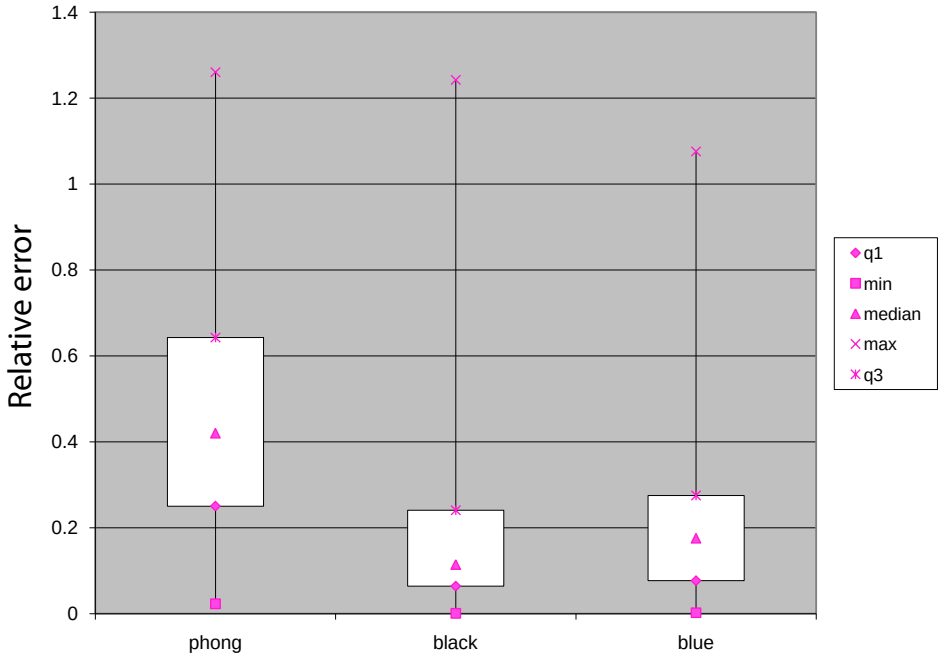


Figure 15: Absolute difference between the estimated and ground-truth relative depth for local illumination with Phong, black and blue shadows. The box plot shows the minimum, maximum, median and the interquartile range (q_1-q_3). Errors greater than 1.0 indicate wrong depth ordering of points.

B.6 Conclusion

We have described a technique which allows for mimicking the illustrative shadowing technique: combining the luminance with the chromatic shift to express shadowiness. This technique allows for better perception of details in the shadowed areas which we proved in two experiments with users. We evaluated surface and contrast perception for a palette of shadow colors ranging from dark gray to light blue. The two experiments confirmed the advantage of our method in terms of perception of detail in shadows. Our experimental setup does not detach the perceptual gain induced by luminance shift from gain induced by the chromaticity shift and so we evaluated the perceptual gain for this particular artistic shadowing technique. Therefore, a further study would be needed to determine how much of the gain is induced by the chromaticity shift itself.

We also studied how our method affects the depth cues from images compared to the original shadowing technique and Phong lighting with no shadowing. We conducted an experiment where subjects estimated relative depth between selected points. The analysis revealed that the depth cue from image becomes slightly degraded as the shadow becomes brighter. Therefore, the chromatic- and luminance shift should be combined with consideration. From the first two experiments, we observed that the most significant

improvement happened after the first chromatic shift from the dark gray shadow. Therefore, we suggest interpolation towards S_1-S_2 as the optimal combination which extracts the best of both methods: depth cues and perception in detail in the shadow.

Acknowledgements

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 and partly within the Geoillustrator research project (# 200512), which is funded by Statoil and the PETROMAKS programme of The Research Council of Norway. The authors wish express special thanks to the certified medical illustrator Kari Toverud for consulting, Helwig Hauser and Maneesh Agrawala for their comments and advice, numerous respondents for participating at the user study, and anonymous reviewers for their comments.

PAPER

C

A Perceptual Statistics Shading Model*

Veronika Šoltészová[1], Çağatay Turkay[1], Mark Price[2], Ivan Viola[1,3]

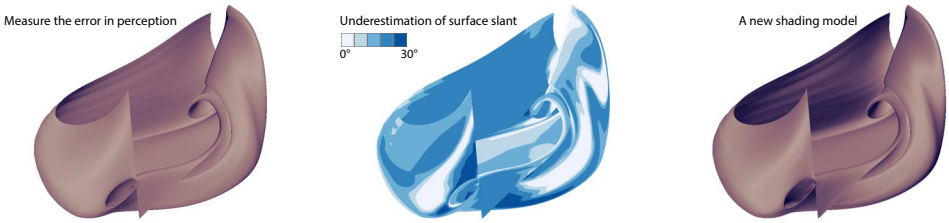


Figure 1: The concept of iterative evaluation-analysis-redesign of a visualization technique is shown on a stream surface dataset. Analysis of the perceived surface slant while using a chosen shading model – the Lambertian shading model [43] on the left leads to a statistical model of the perceived error 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 the 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 which 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.

* This article was accepted to *IEEE Transactions on Visualization and Computer Graphics* and in October 2012, it will be presented at IEEE VisWeek in Seattle, U.S.A. by V. Šoltészová

[1] Department of Informatics, University of Bergen, Norway
[2] Faculty of Psychology, University of Bergen, Norway
[3] Christian Michelsen Research, Norway

C.1 Introduction

The major effort of computer graphics initially focused on the production of synthetic scenes which 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. 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. These local properties include, e.g., patterns of reflected light which are based on the surface orientation and the illumination direction, and texture deformation which is based on local curvature of the surfaces.

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 [51, 140]. Some rendering styles abstract from the realistic scene appearance by exaggerating the Lambertian shading gradient transitions [132]. Even though this approach has initially mimicked artwork, an increasing number of techniques are now motivated by new knowledge from vision research [153, 154]. 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 [60, 118].

The shading models mentioned above have an *imperative* character – an algorithm *dictates* the visual appearance which 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 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 orientations from shading, utilizing a common shading model (Figure 1 left). We perform statistical analysis on data collected from a perceptual study that reveals systematic errors of human visual 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 according to the scene characteristics. This knowledge is captured by datasets that will become publicly available at the time of manuscript publication,
- a new shading model: we obtain a new shading model from the iterative evaluation and improvement concept that enhances surface 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, they have not used 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.

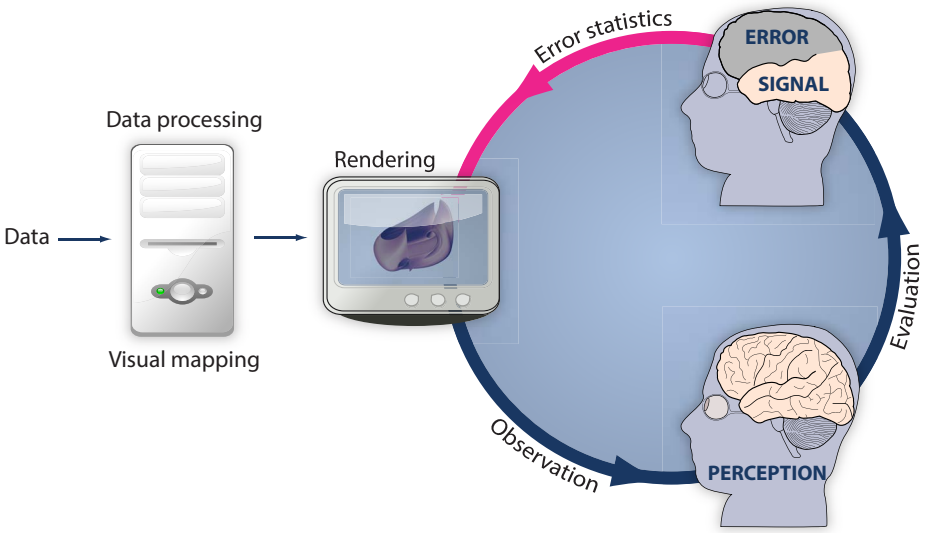


Figure 2: The concept of iterative evaluation and design of a rendering technique. The original visualization pipeline contains no cycles and ends at the stage when the image is perceived by the user. The new concept contains a loop; The accuracy of perception is evaluated and the original rendering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

C.2 Previous Work

For two millenia, scientists have been trying to elucidate the mechanisms in the human visual system (HVS) which 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 [147].

C.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* [120].

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) [147]. 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 [23]. However, shading alone cannot yield the depth structure of a scene correctly [33]. The depth cues from shading are poor when compared to the retinal disparity (stereopsis) and kinetic cues [62].

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. [6] introduced the term *bas-relief ambiguity*; when an unknown object with Lambertian reflectance is viewed orthographically, there 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 evidence that the *pictorial relief*, i.e., 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 [33, 147]. 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 [147].

Mamassian and Kersten investigated the perception of local surface orientation on a simple smooth object, under various illumination conditions [87]. 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

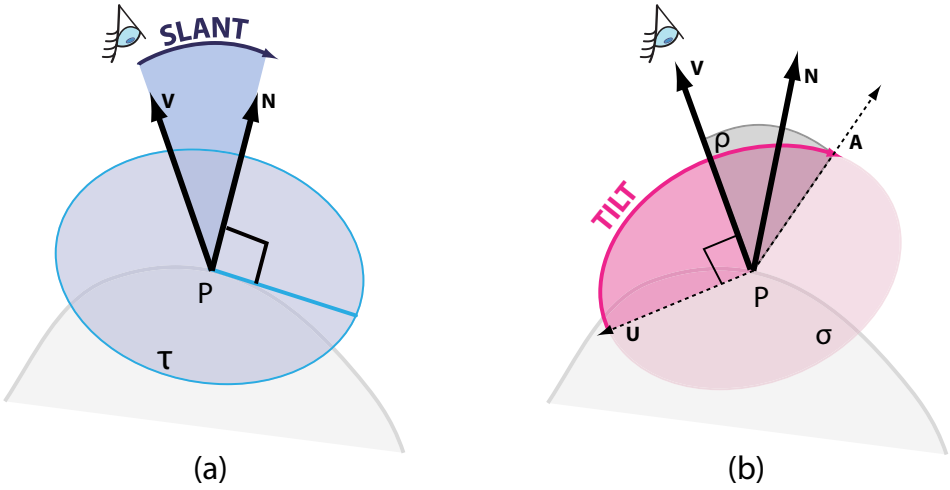


Figure 3: A slant angle θ is defined as the angle between the surface normal **N** at a point P and the viewing vector **V**. τ denotes the tangent plane at P and **U** the up vector of the viewer's coordinate-system. σ is a plane such that $P \in \sigma$ and $\mathbf{V} \perp \sigma$ and ρ denotes the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

the eye space [23]. 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 [47]. Van Doorn and Koenderink [151] suggested that it is the presence of a frame that tends to flatten the pictorial relief in a systematic fashion.

To resolve these ambiguities, the HVS tends to assume a certain light direction. Johnston and Passmore suggested that the slant discrimination declined with rotation of the light direction vector towards the viewpoint [62]. Follow-up studies indicated that this direction is from above the viewer and 12° left from the vertical axis [86, 145]. O'Shea and colleagues studied the assumed slant of the light direction on purely diffuse surfaces with no shadows [109]. They demonstrated that the surface slants were most accurate when the light source was $20° - 30°$ above the viewer.

Fleming et al. studied mirror-material surfaces, i.e., surfaces riddled with specular highlights that contained no shading [40]. They concluded that the HVS can somehow exploit specular reflections to recover three-dimensional 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 [105, 148].

In illustration, artists tended to exaggerate salient features such as curvature or important lines. Their methods have been mimicked by the graphics community. Exaggerated shading [132], geometry manipulation [68], light warping [153] and radiosity

scaling [154] 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 which yields a measurable systematic error in perception.

C.2.2 Psychophysical experiments

The first experiments to human investigate 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 [147]. In the experiment of Mingolla and Todd [98], 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. Their conclusions were threefold: (a) The HVS initially assumes that all surfaces have Lambertian reflectance, (b) the illumination direction must be known before shape detection can proceed, (c) surface orientation is detected locally, and global shape is determined by smoothing over local surface orientation.

The works of Koenderink et al. [73] and Todd [147] 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.

Gauge-figure task:

This task, designed by Koenderink et al. [72], allows one to determine the perceived orientation of a surface. It uses a Tissot's indicatrix, i.e., an ellipse of distortion – a mathematical tool which characterizes distortions from a map projection. When the indicatrix is aligned with a surface which 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 placement 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 [109]. Šoltészová et al. utilized this test to compare the surface perception for different styles of shadow rendering [140].

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 [23]. Their experiment is the most relevant for our work. Their study was performed on 14 different images, both

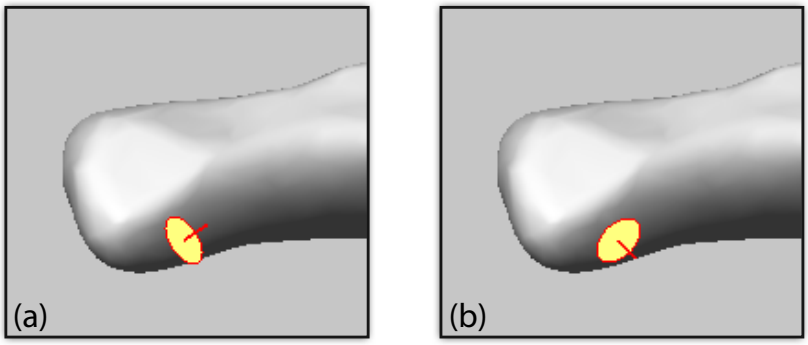


Figure 4: Example of (a) a bad placement and (b) a good placement of a gauge figure (red-yellow Tissot's indicatrix) over a shaded surface.

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 [73]. 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. The 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.

C.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 [33, 147].

C.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 [109, 147]. The slant angle is understood as the angle between the surface normal and the viewing direction. We describe this effect with a mathematical model which we obtained based on 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 [109] or to use a simple shearing transformation in depth [147]. 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 C.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 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 [23]. The overall dependency of estimated surface slant θ_E and the ground truth θ_G slant is approximated with

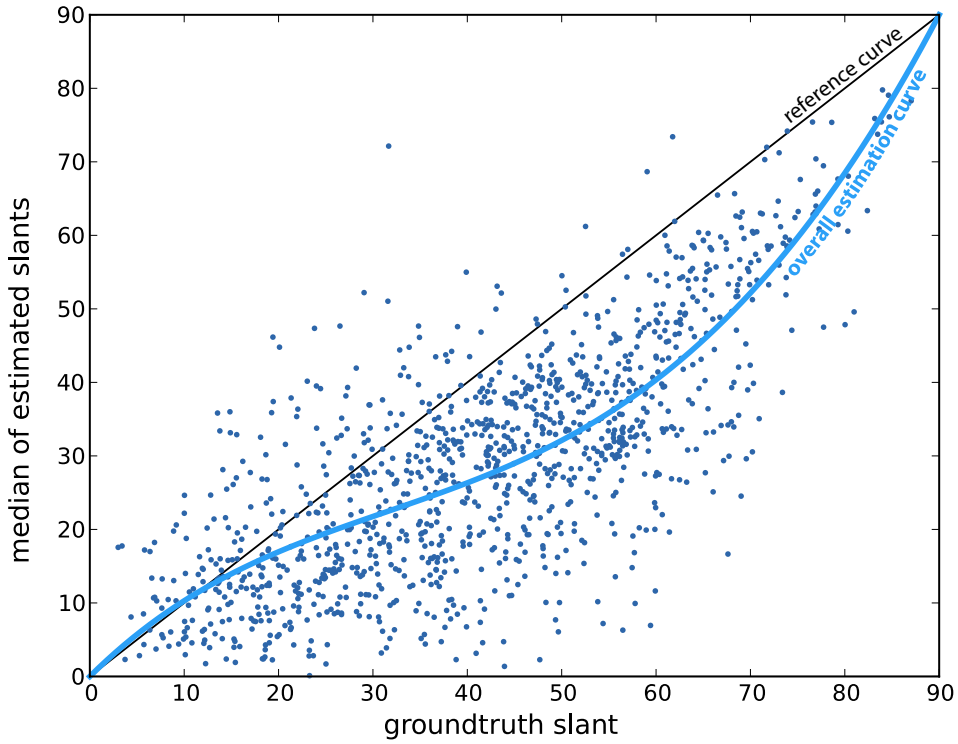


Figure 5: Perceived surface slant as a function of the ground truth slant extracted from the dataset of Cole et al. [23]. Each dot represents the median of the entire set of trials at one sampling position. The *overall estimation curve* is polynomial curve which is fitted to the data. The *reference curve* $x = y$ indicates how a perfectly accurate estimation.

a polynomial fitting curve of the $4^{th} - degree$ 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 [90]. For various types of fit, we obtained the following R^2 values: Fourier fit of 1^{st} degree $R^2 = 0.773$, Fourier fit of 8^{th} degree $R^2 = 0.780$, exponential fit $R^2 = 0.774$, cubic fit $R^2 = 0.773$, and for polynomial fits of 4^{th} degree $R^2 = 0.775$, 5^{th} degree $R^2 = 0.775$ and 8^{th} 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 4^{th} 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 $\phi \in (315°, 45°]$, right or east $\phi \in (45°, 135°]$, downwards or south $\phi \in (135°, 225°]$, and left or west $\phi \in (225°, 315°]$. We define tilt (consistently with the work of Cole et al.) as the azimuth angle on a compass where the wind directions are $N = 0°$, $E = 90°$, $S = 180°$ and $W = 270°$.

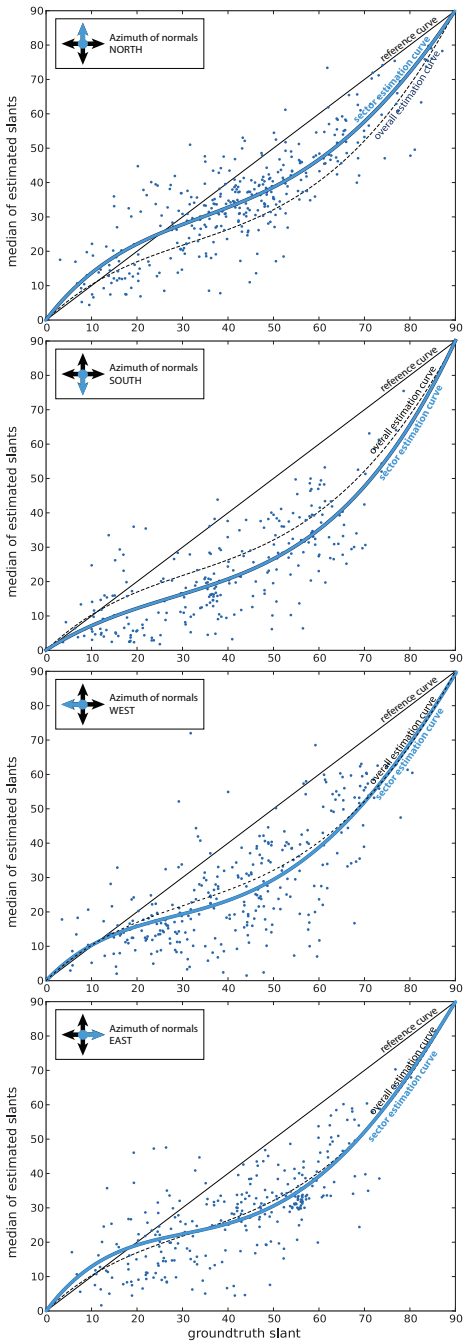


Figure 6: Approximation of slant perception in four different sectors according to the tilt of the ground truth surface normal. The reference curve indicates a perfectly accurate perception while the sector estimation curves approximate the perception of slants in their respective sectors. We also plotted the overall estimation curve which indicates the average perception of slants in all sectors.

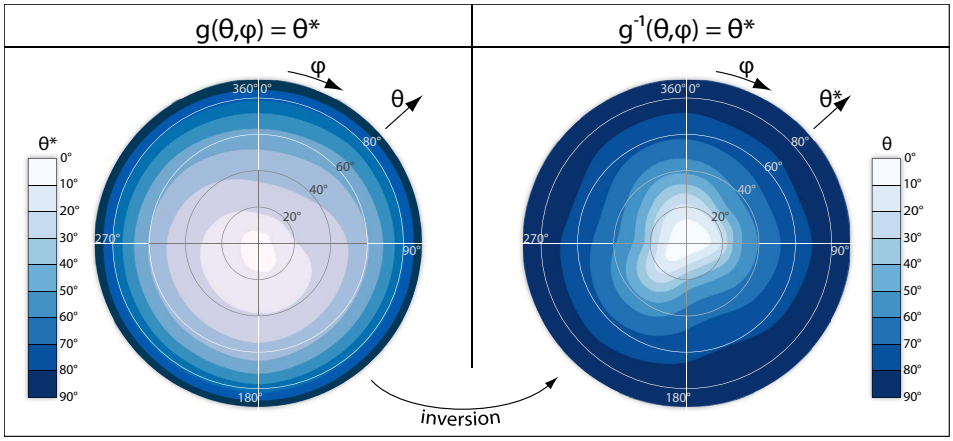


Figure 7: Functions $g(\theta,\phi)=\theta^*$ and $g^{-1}(\theta^*,\phi)=\theta$ rendered as color-coded plots. Since $f=g^{-1}$, the right plot is also the look-up map which allows to efficiently find the slant angle θ of a normal which is perceived as θ^*.

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 [147].

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 expect this threshold to be approximately 20° [87] which is consistent with our finding $15^\circ - 25^\circ$.

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

C.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 $[0°,90°]$. 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 [90]. 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 will be publicly-available for download. 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 $g_{N=0°}^{-1}$ and $g_{S=180°}^{-1}$, yields, even thought not identical, but very similar 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:

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

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.

C.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 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.

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 which are input into our shading model using a perceptual model corresponding to the original shading algorithm. In Section C.3.2, we described how to obtain such 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 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. The components of the modified normal n' are then defined by the following equations:

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

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 perceptual model is applicable to any illumination computation scheme which is based on surface normals or gradients. To demonstrate the effect of our approach, we applied our model to Lamber-

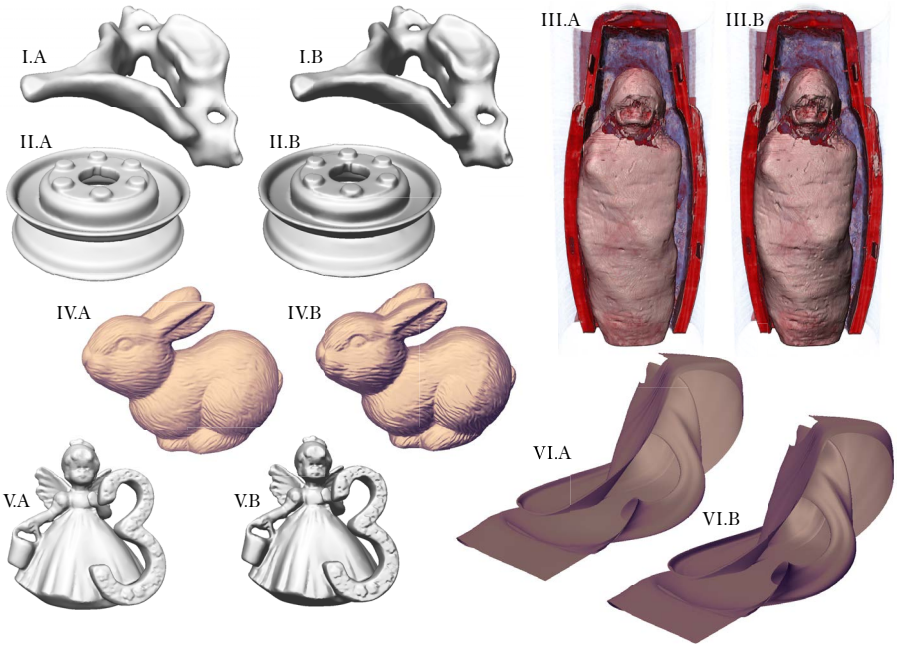


Figure 8: The Lambertian shading using original normals (A) versus statistical shading model (B) shown on various datasets: I – cervical and II – pulley [23], III – a CT scan of a mummy, IV and V are geometry representations of laser scans of a bunny and an angel, and VI – a stream surface.

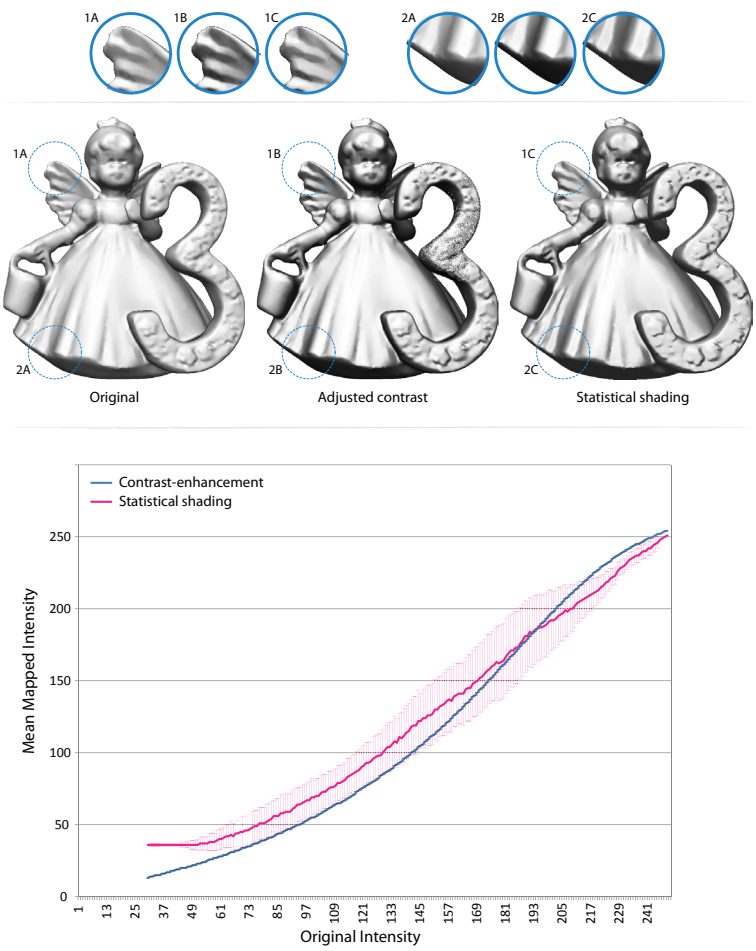


Figure 9: Comparison of a contrast-enhanced image and a statistically-shaded image. We plotted the mean mapped intensities of a contrast-enhanced image and statistical shading. The error bars represent the standard deviation σ of the mapped values.

tian shading and used purely diffuse-reflective materials. In all settings, the light source conforms to the assumed light direction [109]. 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 geometry and as volumes 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 on a range of intensities. In contrast, global contrast enhancement maps the intensities on 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, same results cannot be obtained by simply enhancing the contrast of the image.

C.5 Verification

Our hypothesis here 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 the statistical shading condition (B). We then formally analyzed the difference in performance between the two conditions.

C.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 coworkers [23], 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 [165] and defined their source of illumination as an environment map. We used 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 $[0°, 90°]$ 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 all test cases so, in total, we collected at least 26 samples per test case and more than 8500 solved test cases overall.

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.

C.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 4^{th} 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* $U(\theta_G) = \theta_E$:

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

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).

C.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., $E(0°,20°)$, $E(20°,40°)$, $E(40°,60°)$, and $E(60°,80°)$. The division into subintervals was selected on a priori grounds. According to previous evidence [87] and also concluding from our own analysis, the underestimation of slant is zero at ca. $20°$ of ground-truth slant and highest for slants $40° - 60°$ (see also Figures 5 and 6). Hence we are 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 [46, 92]. 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

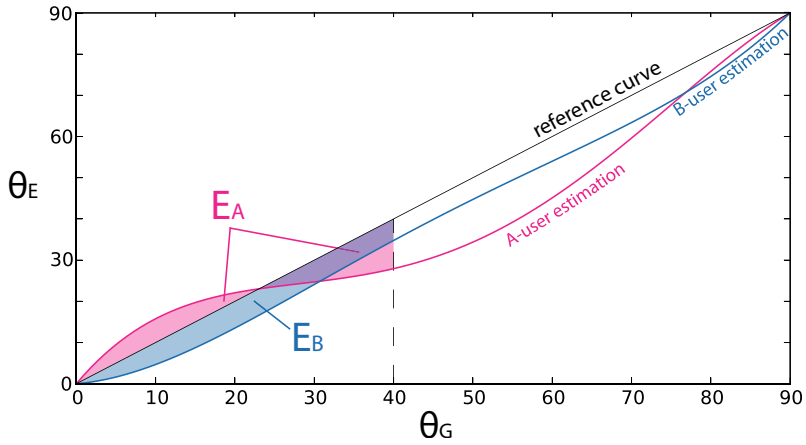


Figure 10: The error areas of a selected participant for the shading condition A – $E_A(0°,40°)$ filled with pink and for the shading condition B – $E_B(0°,40°)$ filled with blue.

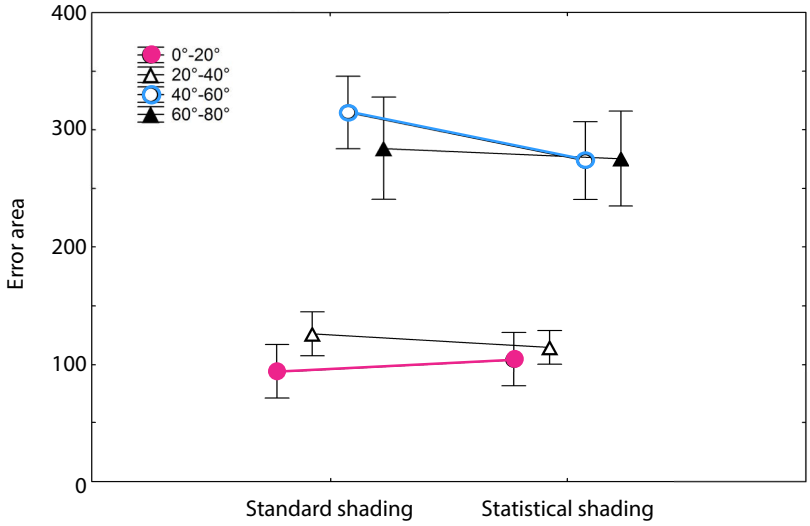


Figure 11: Interaction plot between E for the two shading conditions: standard (A) and new (B) in each of four subintervals of the curve. The vertical bars denote the 0.95 confidence interval. We found a significant improvement in the interval $[40°,60°]$ – blue, a non-significant worsening in $[0°,20°]$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

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 of the algorithm was only obtained for the interval [40°,60°] [F(1,39) = 12.4, p = 0.001, r = 0.49]. For intervals [0°,20°], [20°,40°], and [60°,80°] 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 [160] due to violations of normality for those distributions, 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°,20°], [20°,40°], and [60°,80°] is rather small compared to the interval [40°,60°] 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°,60°]. 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 [20, 21], 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.

C.5.4 Discussion

Based on the results obtained in our gauge-figure experiment, we created a second model of correction as described in Section C.3.2 and applied a second correction. 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 accurate perception of normals slanted $40° - 60°$. Our technique is not photorealistic. One could ask whether this is the case for other techniques which mimick 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 [153], and exaggerated shading [132]. 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 $f(\theta,\phi)$ using function $\tilde{f}$ and precise evaluation using the lookup map.

	$E_A(a,b)$		$E_B(a,b)$	
(a,b)	μ	σ	μ	σ
(0°,90°)	863.82	274.06	813.34	244.78
(0°,20°)	93.95	73.075	104.1862	72.59458
(20°,40°)	126.34	58.52	114.82	44.6
(40°,60°)	314.9	96.6	273.54	104.15
(60°,80°)	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.

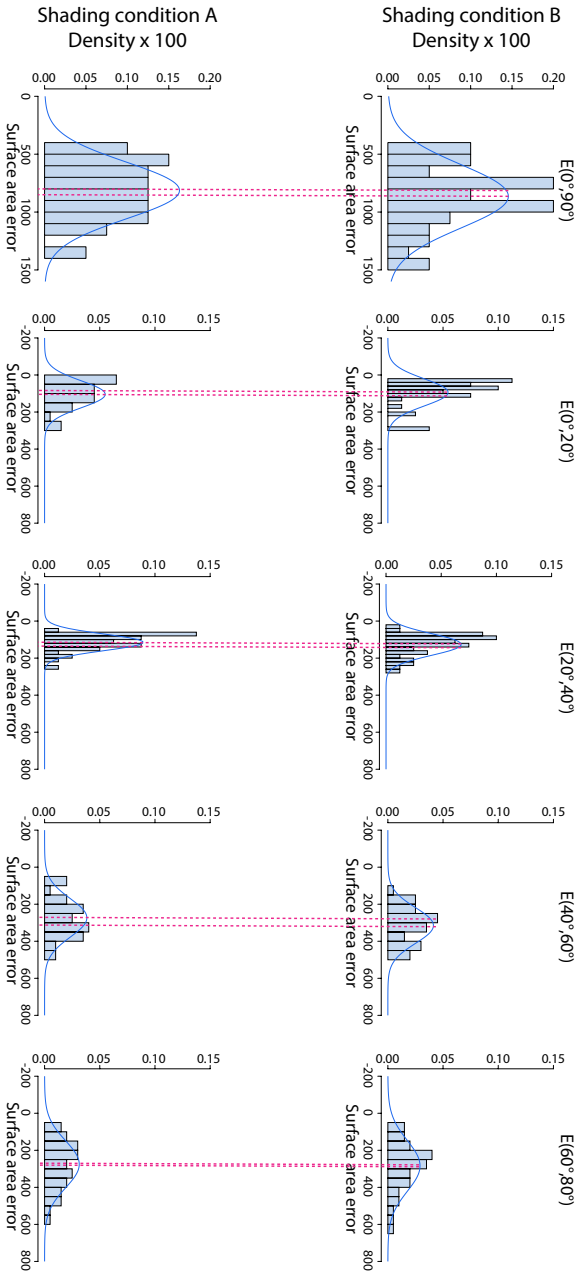


Figure 12: Probability density of the surface area error E with histograms and approximative normal distribution curves for the entire interval $[0°,90°]$ and four subintervals $[0°,20°]$, $[20°,40°]$, $[40°,60°]$, and $[60°,80°]$ and for both shading conditions. In intervals $[0°,20°]$ and $[20°,40°]$, 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.

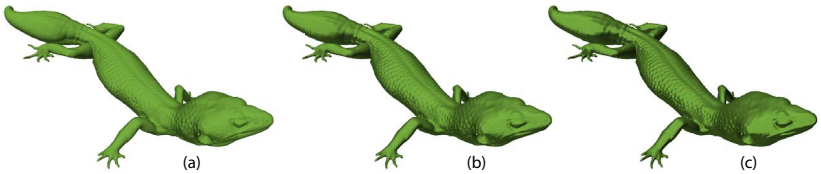


Figure 13: Rendering results of a leopard gecko CT dataset of the iterative process *evaluation and re-design*: (a) the original Lambertian shading, (b) the result of a modification after the first user study, and (c) the result of a modification after the second user study.

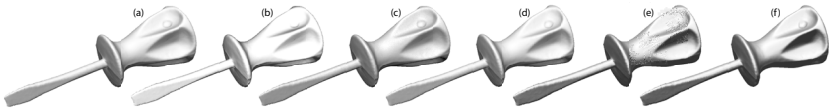


Figure 14: Comparison of (a) the Lambertian shading, (b) light warping, (c) exaggerated shading, (d) shearing along the z-axis and our approach using (e) the approximating function $\tilde{f}$ defined in Equation 1 and (f) the precise look-up texture to solve $f(\theta,\phi,)$.

C.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 θ^* should be presented so that it is perceived as θ. 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 is the highest of the slant perception is the highest.

C.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.

C.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, were to be used, or if the objects were 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 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 [23]. 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.

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 [154]. 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.

C

Perceptual-Statistics Shading

PAPER

D

Stylized Volume Visualization of Streamed Sonar Data*

Veronika Šoltészová[1], Ruben Patel[2], Helwig Hauser[1], Ivan Viola[1]

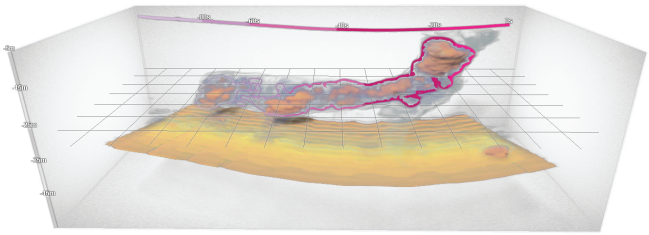


Figure 1: Visualization of a large fish school of sand eel floating above the sea bottom reconstructed live from a series of 2D slices. The temporal outline color-encodes the temporal dimension of the volume visualization.

Abstract

CURRENT visualization technology implemented in the software for 2D sonars used in marine research is limited to slicing whilst volume visualization is only possible as post processing. We designed and implemented a system which allows for instantaneous volume visualization of streamed scans from 2D sonars without prior resampling to a voxel grid. The volume is formed by a set of most recent scans which are being stored. We transform each scan using its associated transformations to the view-space and slice their bounding box by view-aligned planes. Each slicing plane is reconstructed from the underlying scans and directly used for slice-based volume rendering. We integrated a low frequency illumination model which enhances the depth perception of noisy acoustic measurements. While we visualize the 2D data and time as 3D volumes, the temporal dimension is not intuitively communicated. Therefore, we introduce a concept of *temporal outlines*. Our system is a result of an interdisciplinary collaboration between visualization and marine scientists. The application of our system was evaluated by independent domain experts who were not involved in the design process in order to determine real life applicability.

* This article was published in *Proceedings of the Spring Conference on Computer Graphics (SCCG)*, 13–20, 2012 and presented at the named conference in Smolenice castle, Slovakia by Veronika Šoltészová. The paper was awarded by the 2^{nd} best paper award and therefore invited to the *Computer Graphics Forum* journal. Furthermore, the presentation was awarded by the 2^{nd} best presentation award.

[1] University of Bergen Bergen, Norway
[2] Institute of Marine Research, Norway

D.1 Introduction

One of the main goals of marine-fisheries research is to map the processes of marine ecosystems by observations and theoretical work. This includes surveys on research vessels in order to estimate the biomass of the stock and to study processes such as the behavior and morphology of fish schools, *i.e.*, "flocks" of fish.

Scientific instruments on vessels for marine-fisheries research rely heavily on remote acoustic sensing such as 2D and 3D sonars. The visualization software development has stagnated compared to the development of the sonar hardware and only elementary visualization toolkits are currently available; For the 2D sonars, the in-situ visualization is limited to 2D views where the color is a function of the acoustic reflectance. Volume visualization is available only as postprocessing [74, 137]. Basic volume visualization is available only for the 3D sonars [94, 4]. However, 2D sonars have lower cost and can be affordable for fishing vessels. Therefore, dedicated in-situ volume visualization based on an input from a 2D sonar is worth aiming at.

Even though the structures of schools were revealed to be more complex [99, 110], they can be viewed only as compact, homogeneous units. Because of such limitations in visualization technology, the focus of scientific methods in fisheries can be oriented only on the quantitative, non-visual analysis of the data such as rudimentary measurements of the biomass [85]. Nevertheless, the structure and behaviour of fish schools is of high interest because it can help to explain their yet poorly understood ecological meaning.

In collaboration with marine scientists, we designed an application for fast volume rendering intended to be used in-situ on 2D sonars which scan the water column vertically. We fill the gap in the dedicated visualization technology – we propose a visualization tool which addresses the needs of the marine-research domain and brings the following contributions to the state-of-the-art in visualization:

1. A successful application of 3D visualization based on 2D scans and time for a new scientific domain.

2. An innovative architecture for an efficient volume rendering method which operates directly on the 2D scans without prior resampling the space on a voxel grid. Thus, we reduce the number of resampling stages.

3. A new time-to-live concept and an efficient storage mechanism tailored for the in-situ visualization of streamed images.

4. A novel concept called *temporal outlines* as shown in Figure 1. They clearly communicate the temporal nature of the volume visualization – their color and thickness associates parts of the volume with the time of acquisition.

Based on the discussions with independent domain scientists, we describe the expected use of our system. Our application extends the utility of 2D sonars and is an evolutionary step in the visualization technology used in the marine domain.

D.2 Related work

Volume visualization techniques have been investigated for more than two decades. Sophisticated algorithms are now available and allow for visualization of large data in real-time and offer excellent visual enhancements. However, only a few techniques are tailored to 3D visualization of sonar imaging. For example, a framework called *SonarExplorer* serves for the analysis and visualization of fish schools tailored to 3D-sonar imaging [4]. Explicit imaging of sonar data in 3D as in *SonarExplorer* can be achieved only using acquisitions from a 3D-sonar. These are highly expensive and therefore, 2D-sonars became a commodity.

Visualization of ultrasound images (US) in medicine is facing a similar problem: 3D probes exist but 2D probes are cheaper and more widespread and in addition, yield images of superior resolution. Without a 3D probe, doctors can acquire a volumes using freehand US systems: a 2D probe with an attached positional-tracking system which allows for volume reconstruction as post-processing. The acquisition by a 2D-sonar is analogous to the freehand US in medicine. A vessel equipped with a 2D-sonar is equipped with a high-end GPS device, a motion reference unit and a clock so that every scan can be positioned in space and time. Eventually, the volume can be reconstructed in a post-processing step [137].

Volume rendering of data consisting of a sweep of 2D scans requires prior reconstruction to fill the gaps between the scans on a regular voxel-array. In the state-of-the-art technique on volume reconstruction from freehand US [65], an optimal orientation of reconstruction slices is selected. The volume is reconstructed slice-by-slice while following this direction: The scans are sorted according to the acquisition time and successors are paired. The intersection lines of these pairs and the current reconstruction plane define a polygon. The intensity values given along the intersection lines are interpolated linearly across the polygon which is then drawn into the reconstruction plane. Volume reconstruction causes a delay and introduces errors at two stages; first during volume reconstruction and second during rendering.

Direct reconstruction during rendering of surfaces from freehand US in medicine was supported by the *stradx* system [121]. *Stradx* included visualization techniques such as slicing where a naive slice reconstruction would extract values along the lines of intersection of each ultrasound scan with the selected slice plane. To improve the quality, they suggested the following interpolation scheme. They did not consider each scan to be infinitely thin but assign it a certain thickness. The intersections of US scans and the slice plane became smeared polygons which overlap over each other. In the overlapping regions, they take the value which is closer to the center line of the respective slab. Later, the system was extended by volume rendering [122]. The slicing method was used for slice-based volume rendering. They stated that the slice reconstruction was fast because as an optimized sequential algorithm was used but no detailed description of their implementation nor any references were given.

Many works address the problems related to the volume visualization of time-varying data. We focus on related work solving data intermixing issues in the context of time-varying volume data [15]. Woodring et al. presented a method for viewing high-dimensional data using a projection on a hyperplane [164]. During the hyperplane projection, they combined information from subsamples using schemes such as alpha composition, first

hit, addition, MIP, average and deviation.

It is not straight-forward to depict motion in a motionless image. Illustrators or comic artists use stylized motion-lines and arrows to deploy motion in static pictures [95]. In non-photorealistic rendering, these methods are often mimicked. Stylized lines can unambiguously create the sense of direction by their thickness and color variation relative to the background. The thinner and sharper end of the stroke with less contrast to the background shows where the flow is coming from [150, 157]. In visualization, stylized temporal gradients and lines were employed to show the movement of time-varying data within one frame [143, 64].

Our concept builds on the previous work on reconstruction [65] and direct rendering of arbitrary scans [122]. Fused reconstruction and rendering of streamed sonar data is a labor-intensive task. We describe an efficient storage system and an architecture designed to harness the power of modern graphics hardware to achieve interactive frame rates. Unlike the previous approaches, our framework integrates a shading model into the rendering pipeline which notably enhances the quality of the visualization. In addition, we provide novel illustrative overlays, *temporal outlines*, which help to depict objects from the background and intuitively communicate the temporal dimension. The application design addresses needs of marine-fisheries research domain and the framework clearly fills the gap in visualization of streamed 2D-sonar data.

D.3 Pipeline overview

Between acquisition and the final visual output, the data passes multiple stages. Figure 2 illustrates the overview of the system pipeline: acquisition, preprocessing and slice view were part of the original package coming with the sonar. We add volume visualization which operates on the images and the positioning information supplied by preprocessing stage. This section leads through individual stages of the pipeline in order to present the overall concept of our system.

Acquisition:

The source of the data is a 2D-sonar of type Simrad *ME70* which is a multibeam scientific echosounder used for biomass estimation, fish school characterization and behavior studies [137]. In our case the sonar transducer is oriented downwards, therefore the water column is sampled vertically as it is illustrated in Figure 2. The data is collected and saved to the disk in a specific format. For each measurement a fan is sampled perpendicularly to the bow of the ship. A fan constitutes data collected during one sampling cycle of the sonar, also denoted as *ping*. The fans consist of 45 electronically stabilized beams with an opening angle of 2° covering a swat of 140°. The frequency band of the sound waves is 70-120 kHz.

During a survey, the sonar samples data continuously at different sampling intervals. The interval is decided by the user, sonar software and limited by hardware. For instance if the user decides to increase the sampling range, the wave propagation time will increase which again can increase the number of data points sampled for each beam. This will increase the total time to sample one fan which in turn increases the sampling time between fans.

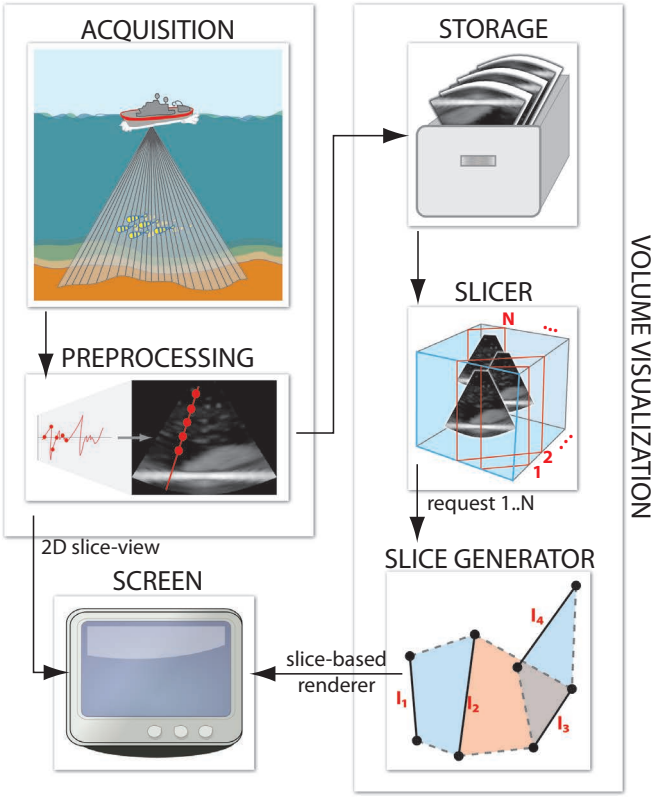


Figure 2: A conceptual overview of the system pipeline.

Preprocessing:

Data from the sonar is converted from waves to beams of acoustic reflection, i.e., fans, and preprocessed to remove noise, mostly on a per-beam basis. Further reformatting of the processed data involves the conversion of individual fans to bitmaps and a header file. The header file contains the transformation matrix and timestamp of the fan. The transformation is captured by a high-end GPS device and a motion reference unit (MRU) which is capable of measuring pitch, roll and heave of the vessel with a high level of precision. The transformation matrix is used by the visualization application to transform each bitmap into world coordinates. This stage is regarded as instantaneous in time regardless of wave propagation time. With these settings and equipment, we retrieved on average 1-2 bitmaps and headers per second.

Volume visualization:

Previously, the only instant visualization stage in the 2D-sonar imaging system was elementary slicing in one dimension. In contrast, our system allows for instantaneous 3D visualization. As the scans are streamed continuously, it is necessary to involve a *stor-*

age management unit. We store the last n scans and their transformations. The renderer operates on the data in the storage system. For each frame, the scans are transformed to the world coordinates using the transformation in the associated headers and then to the viewing coordinates. In the *slicer* stage, the bounding box of the transformed scans is sliced with view-aligned planes as for slice-based volume rendering. We are not operating on a voxel grid and therefore slicing of the bounding box will show only intersecting lines of the slicing plane and scans. Therefore, each slice must be reconstructed in the *slice-generator* stage before it can be rendered and blended with the 3D visualization. The rendering includes optional visual enhancements, in our case shadowing, *temporal outlines* and a reference grid.

D.4 Direct volume reconstruction for rendering

The stream renderer operates on a set of 2D bitmaps which are placed in 3D space and have an associated timestamp. First, we store a set of n most recent scans, where the cardinality n is set by the user. We transform the scans by their corresponding transformation matrices to the world space and fit a bounding box around them. The bounding box represents the proxy geometry for the volume rendering. As the volume defined by the bounding box is not represented by a regular grid, it is not straight forward to design a rendering algorithm for a volumetric dataset with such representation. Contrarily, the data is in our case represented by a set of 2D scans which are arbitrarily placed in 3D space. Therefore, the rendering stage must address an additional reconstruction problem which fills the gaps between individual images.

Previous work on reconstruction from freehand ultrasound in medicine delivers high-quality results but causes an additional delay before the rendering step. This is feasible in some scenarios in the medical domain when the doctor makes his or her acquisition, and inspects the volume after a delay. This is inapplicable for volume visualization of data being streamed-in with frequency 1-2 images per second. Another disadvantage of this approach is that the data is resampled at two stages; during reconstruction and during rendering.

We are fusing reconstruction and rendering by extracting the best from the previous work: reconstructing the volume on a view-aligned stack of planes using a high-quality plane-reconstruction scheme [65] and using reconstructed slices for slice-based rendering [122]. The computation is adapted to exploit modern graphics hardware. We describe a fast volume rendering solution for the stream input of 2D images and provide a novel concept of showing the temporal aspect of the data. While our design choices address the specific needs of the marine domain, the general concept of our system can be applied to other domains, e.g., freehand ultrasound in medicine.

D.4.1 Plane reconstruction

As we showed in Figure 2, the bounding box of scans in the world space is sliced by view-aligned planes. The scans and the slicing planes yield a set of intersection lines as shown in Figure 2 in the *slice generator*. Drawing the intersection line segments only is insufficient and therefore, the *slice generator* fills the gaps between the intersection lines. In Figure 3, we see the scans as textured quads and quad-plane intersections become tex-

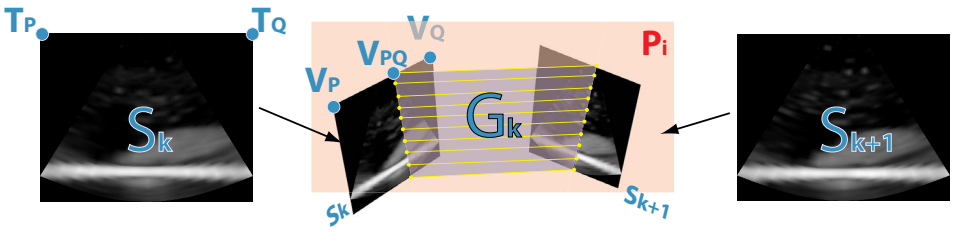


Figure 3: We define a hexahedron by connecting two scans S_k and S_{k+1}. These two faces of the hexahedron are textured. For example, vertices V_P and V_Q have texture coordinates T_P and T_Q respectively. The intersection by a view-aligned plane P_i defines a polygon G_k which patches the gap between the intersection lines.

tured lines. A fast and high-quality interpolation and composition scheme is now needed to fill the gaps between each pair of lines and treat the overlapping gaps.

The storage management keeps the quads in the order they were streamed in, i.e., sorted by their associated timestamps. We connect each pair of successive quads S_k and S_{k+1} in the sorted storage into a hexahedron as illustrated in Figure 3. The intersection of the faces of the hexahedron and a slicing plane P_i defines a polygon G_k which fills the gaps between the intersection lines. For example, the edge of the hexahedron V_PV_Q has texture coordinates (T_P, T_Q) and intersects the slicing plane P_i. The intersection V_{PQ} is defined by a linear combination of V_P and V_Q:

$$V_{PQ} = \lambda_{PQ}V_P + (1-\lambda_{PQ})V_Q \quad with \quad \lambda_{PQ} \in [0,1] \tag{1}$$

The texture coordinates of the intersection T_{PQ} will be calculated as linear combination of T_P and T_Q using the same λ_{PQ} as in Equation 1:

$$T_{PQ} = \lambda_{PQ}T_P + (1-\lambda_{PQ})T_Q \tag{2}$$

It can happen that two successive quads intersect and the connection of their vertices does not yield a regular hexahedron, for example, if the curvature radius of the vessel trajectory is smaller than the extent of the scan in world space. In this case, we split the hexahedron along this intersection into two prisms and treat them as hexahedra with one collapsed edge.

We interpret the bitmap intensity values and time as the first and second color channels in a texture. This allows for smooth interpolation of two attributes: intensities and timestamps (time-to-live). According to the interpolated intensities, we apply a color and opacity transfer function and render the polygon into the current slice. In the case that two polygons overlap, we account for two compositing schemes.

Overwrite:

We replace fragments which have a lower timestamp with the values of the current polygon. The hexahedra are processed in time-ascending order, also the polygons are drawn in time-sorted order. Therefore, the buffer can always be overwritten in overlapping regions.

Alpha blending:

Fragments which have a lower timestamp C_{old} are behind the fragments which have a higher timestamp C_{new}. We are using an over-operator as blending equation for the resulting color C:

$$C = \alpha_{new} C_{new} + (1 - \alpha_{new}) C_{old} \tag{3}$$

D.4.2 Data storage

The scans are stored in a 3D texture similarly as cards in an "index card". When the storage is full, we use the *cyclic queue* principle. The entry position E contains the r-texture coordinate, of the oldest scan stored, i.e., with the lowest timestamp. When a new scan is streamed-in, we render it in the place of the oldest scan and update the entry position. It is convenient to have all scans stored in one texture because only one texture needs to be bound per frame. A scan S_i is then accessed with 3D texture coordinates (s,t,r). The r-coordinate is constant for one scan. Texture coordinates of the intersections are generated using Equation 2.

When calculating the intersections and their respective texture coordinates, we rely on hardware interpolation between scan-pairs in the 3D texture. We create pairs as, e.g., in Figure 4a. As long as the starting timestamp E points to the beginning to the texture, interpolation between pairs according to Equation 2 works correctly. However, if E points another timestamp, two successive timestamps (9^{th} and 10^{th}) would not be neighbors in the 3D texture if we had not one duplicated layer. This parity $(n+1)^{th}$ image is a duplicate of the 1^{st} image and ascertains that the interpolation of texture coordinates works correctly.

Our interpolation scheme delivers perfectly correct results for straight vessels trajectories. In order to obtain correct results for any type of trajectory, we would need to account for rotation in our interpolation scheme for a substantial performance penalty. Coupé et al. described a solution for this problem in the context of post-processing volume reconstruction from freehand ultrasound in medicine [27]. They also analyzed the reconstruction error of their technique, nearest-neighbor interpolation and linear interpolation such as ours. It is also worthwhile noting that their improvements in the reconstruction quality are significant only for sparse scans and large rotation angles. Extending our interpolation scheme by the idea of Coupé et al. would be a potential improvement of the precision but according to our collaborators, is currently not needed.

To further increase the performance, we also allow for scan-skipping: the pairs are created from every k^{th} scan. Figure 4b illustrated pairing for scan skipping $k = 2$. In order to close the chain, the last scan is always paired even though its pair skips $< k$ scans. In Figure 4b, the chain is closed with timestamps (9,10) and (11,12) even though $k = 2$. The skipped scans are not removed from the storage. As they are located physically between the paired scans, and therefore, the interpolated texture coordinates will point to the "skipped" 3D-texture space. Scan-skipping with hardware interpolation over the skipped frames is possible only if we involve 3D texture in the storage mechanism and not a vector of 2D textures used in the previous work [65].

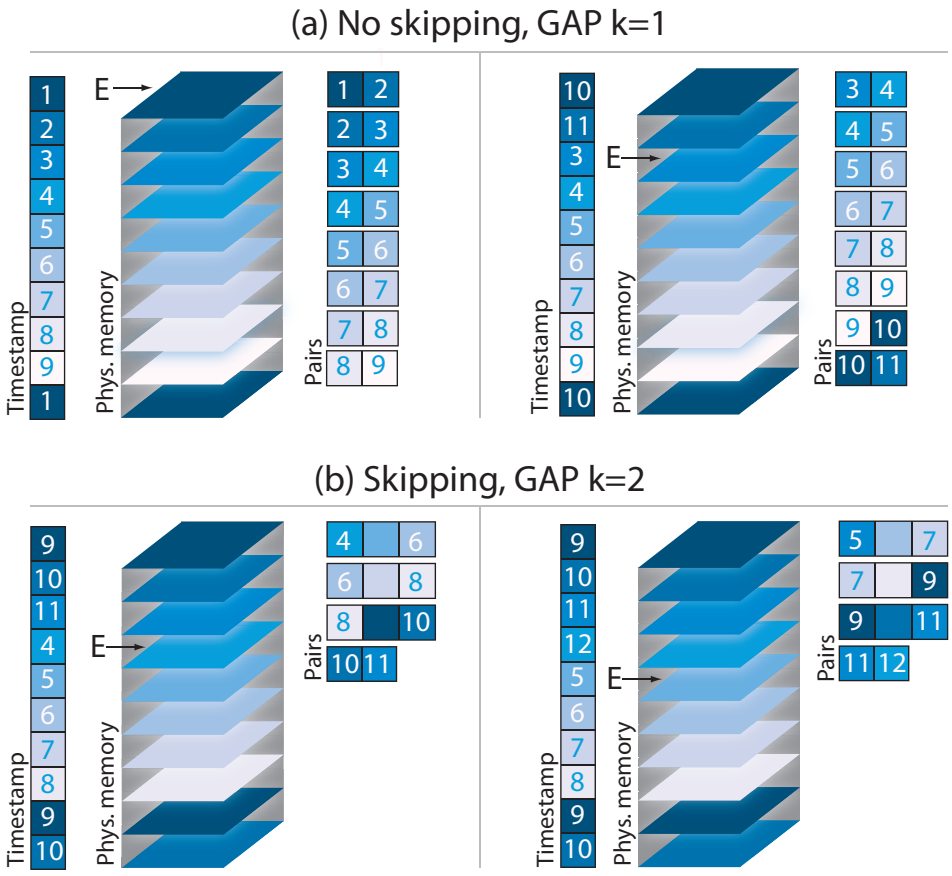


Figure 4: The *cyclic queue* principle of our storage system. Examples of how scans located in physical memory are paired according to their timestamps without (a) and with (b) skipping.

The skipping reduces the overhead connected with the calculation of the intersection geometry but is a trade-off between precision and performance and the k-factor can be used to achieve better performance. The skipping-induced error depends on the relative displacement in space between successive scans related to the vessel trajectory but can be controlled by the application. Three cases are illustrated in Figure 5. The exact skipping-induced error estimation is beyond the scope of this paper but would allow for automatic selection of the k-factor with respect to a preset error tolerance.

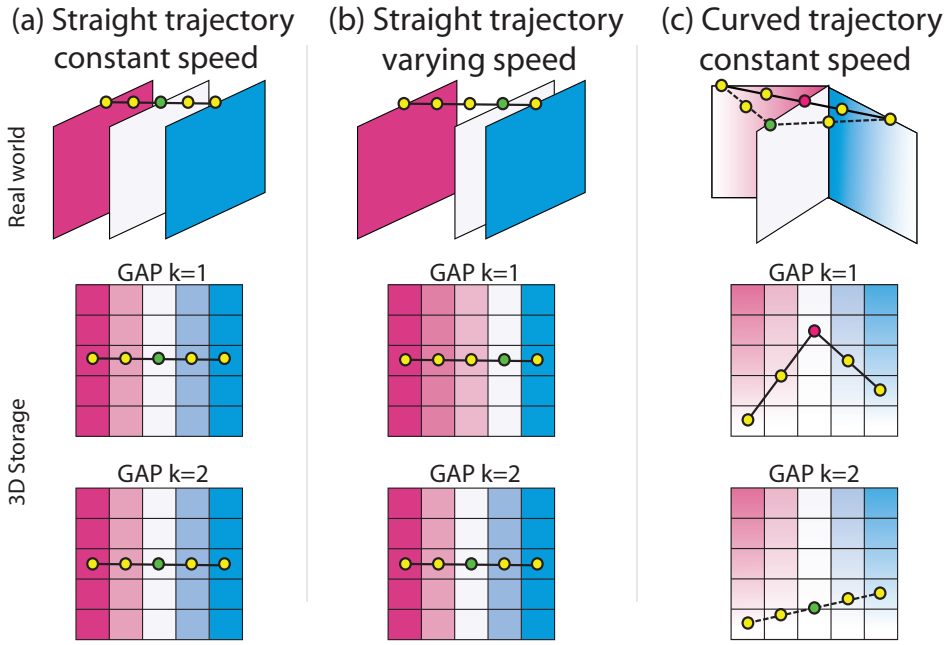


Figure 5: Errors induced by scan-skipping. (a) No error when the vessel follows a line at constant speed, all samples have correct texture coordinates. Error in interpolated texture coordinates when (b) it changes speed, and (c) when its speed is constant but the trajectory is curved.

D.5 Temporal outlines

According to an independent domain expert from Simrad, the hardware vendor, *time is very important as this is a dynamic situation. The biology constantly changes as a function of time. It is absolutely necessary to know what is new and old information in order to know the actual situation and how things have changed, e.g., in which direction is the biology moving and how does the behavior change*. We introduce a new concept of illustrative outlines which convey the temporal dimension of our data which is otherwise not communicated. In our case, users observe a non-static environment and temporal outlines are features which allow to establish a degree-of-confidence in different parts of the visualization. The temporal dimension is encoded by their thickness and color of the outline. A temporal outline is an image-space effect which is calculated after the volume rendering stage as an image-space effect. Therefore the performance consumption is negligible compared to the volume rendering process. The concept is easily applicable for other volume rendering techniques with a temporal attribute.

The generation of time-dependent outlines happens in three steps.

1. Buffer generation: We create a binary image segmentation, i.e., a *membership buffer*, where each pixel belongs either to the fish school or to the background and a time-stamp buffer. The time-stamp buffer stores the time attribute of pixels which are members of the object in the membership mask.

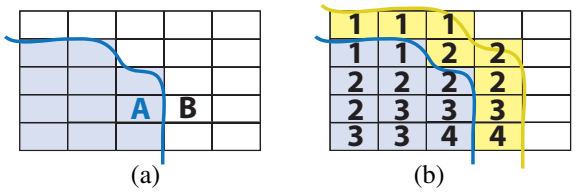


Figure 6: A membership buffer (a) and a time buffer (b). Pixels A belong to the object and B to the background. Pixels B contain no timestamp and therefore, the stamps from the border of A are propagated into the narrow band shown in yellow.

2. Dilation and blurring: We blur the mask with a kernel which size is time-dependent – pixels with a low timestamp (more recent) are blurred more than pixels with a high time-stamp (older).

3. Rendering: We render the contours using the difference image between the blurred and original mask similarly to unsharp masking [84].

Buffer generation:

A rough separation of fish school from the water and the sea bottom can be achieved by thresholding. Fish schools yield stronger echo than water but weaker than the sea bottom. During volume slicing and reconstruction, we threshold the sampled values against a user-defined threshold. Different threshold values allow the user to quickly explore the density levels (MIP-levels) of the fish school. If at least one value satisfying the membership condition is sampled along a viewing ray, the pixel in the membership buffer $\mathscr{M}(s,t)$ attributed to this viewing ray will be 1 and 0 otherwise. A pixel of the time buffer $\mathscr{T}(s,t)$ contains the highest timestamp along the corresponding viewing ray if the pixel is a member of the object and -1 otherwise. The timestamps are originally integer values $[0,n]$ where n represents the current time but later on scaled to the interval $[0,1]$.

Dilation and time-dependent blurring:

The size of the Gaussian convolution kernel $\mathscr{K}$ we use for blurring the membership buffer in pixel $\mathscr{M}(s,t)$ depends on the timestamp which is stored in the time buffer $\mathscr{T}(s,t)$. Figure 6a shows an example of $\mathscr{M}$. After the blurring operation, the values from pixel A in the mask should have "spread" to pixels of the background B. However, pixels of B contain no timestamps and consequently, the kernel size in B would be undefined. Therefore, prior to the blurring operation, we perform a dilation step on time buffer in order to spread valid timestamps into a narrow band around A. The dilation is illustrated in Figure 6b. The width of the band is defined as the largest possible radius of $\mathscr{K}$. According to our propagation rule, we take the closest timestamp. If more valid equally-distant stamps are found, we take the more recent one (higher timestamp). After the propagation, $\mathscr{M}$ is blurred with a Gaussian kernel $\mathscr{K}$ of size $f(\mathscr{T}(s,t))$:

$$f(\mathscr{T}(s,t)) = 8(\mathscr{T}(s,t)) \quad with \quad \mathscr{T}(s,t) \in [0,1] \tag{4}$$

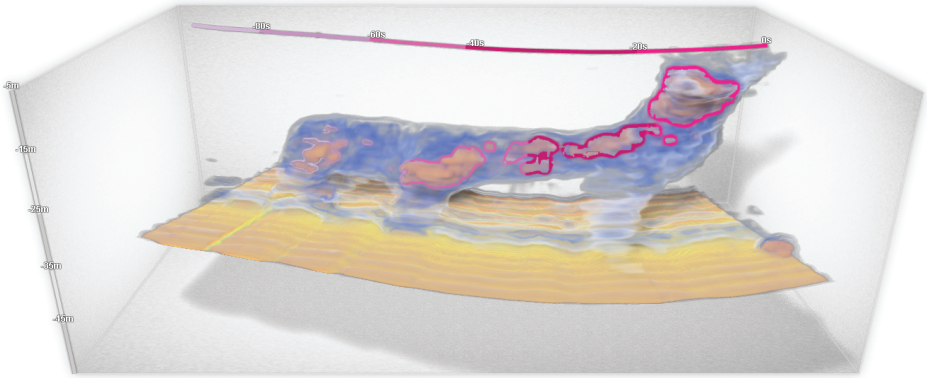


Figure 7: Integrated quantitative measurements: curved time axis above the fish school and the depth axis on the left. We also integrated an optional horizontal grid which is shown in Figure 1.

The outlines are calculated in the image-space and their thickness is independent of the distance to the viewer. Outlines attributed to more recent parts of the volume are thicker and thus more prominent.

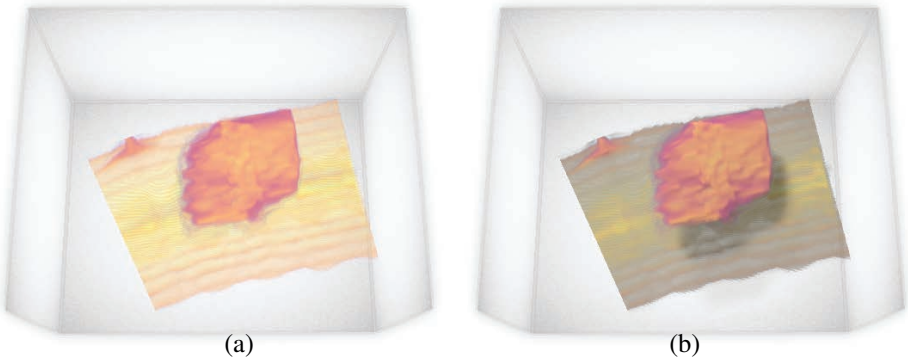
Outline rendering:

The outline is defined as a set of pixels at non-zero differences between the original and blurred membership buffer. If this difference is non-zero, we perform a color look-up based on sampled value from the underlying time buffer. The color is fetched from a 1D color transfer function. The color transfer function is a texture which contains low number of color bands, e.g., 5-6. Each color encodes one time interval. The color with the highest contrast to the background shows the most recent time interval as in Figure 1. The color quantization to a low number of colors in the transfer function helps to read the time interval because it is more difficult to read the time from a continuous map. Finally, the outlines are drawn as an overlay over the volume visualization.

The construction of temporal outlines is similar to rendering of halos using unsharp masking of the depth buffer [84]. The difference to our approach is that we are performing blurring of the membership buffer instead of unsharp masking the depth buffer and most importantly, our blurring kernel has not a constant size.

Integrating quantitative measures:

The colored temporal outlines require a legend showing which time interval corresponds to which color. Instead of showing the legend next to the visualization, we embed it in the volume view. The legend appears as a colored curve which follows the trajectory of the sonar-transducer as shown in Figures 1 and 7. The labels of the time curve appear at the boundaries between individual color bands and show the elapsed time in seconds from the moment when the transducer was located at that point. As the legend is integrated and



(a) (b)

Figure 8: Visualization of floating fish school with no illumination (a) and with applied multidirection shading model (b). Shadows resolve the ambiguous depth cue – in the visualization without shadows, it is not clear whether the fish school is attached to the bottom.

not drawn on the side, users efficiently associate time with parts of the vessel trajectory and parts of the visualization. The curved legend represents the time axis, but the users also need to assess the spatial dimensions in order to grasp the extent of features in the fish schools. Therefore, we integrate also the depth axis which is snapped to one of the horizontal edges of the bounding box and an optional grid. The distance between individual grid lines is fixed to 10m.

D.6 Implementation details

We implemented a proof of concept of the described method using C++, OpenGL/GLSL into the VolumeShop [10] framework. In this section, we describe an important optimization of the intersection generating stage and display the performance of our implementation. Finally, we theoretically describe requirements for a raycaster which operates on the same input data and reconstructs the volume during raycasting in fragment shaders and shows why such implementation is infeasible.

Generation of intersections:

The calculation of geometry can easily become a bottleneck in our pipeline if not designed carefully. The slice generator needs to produce intersecting polygons for each slice plane. This implies a complexity of $O(mn)$ with m number of slicing planes and n number of hexahedra. It is therefore worthwhile to parallelize this process. Our system uses fast hardware-based geometry generator using geometry shaders. The input geometry for the shader consists of the individual hexahedra. As the input geometry is the same for each slicing plane, we define it as a vertex buffer object (VBO) and upload it to the graphics processor only once per frame. The VBO is updated with every ping when a new scan is produced, i.e., maximum twice per second. The intersecting polygons present a

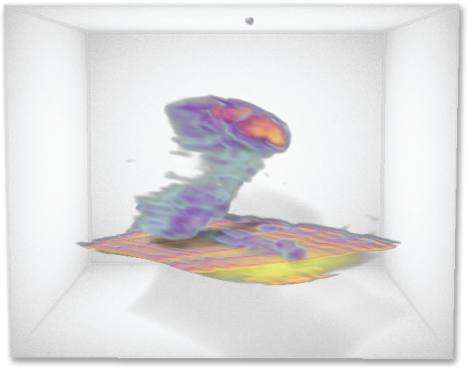


Figure 9: Sand eel fish often school towards the sea bottom. In this case, there is only a relatively small gap which is revealed by the shadow cast on the sea bottom.

substantially larger amount of geometry then of the hexahedra and are in addition viewer-dependent. The hexahedra-definition in the VBO is view independent. It is therefore advantageous to send a smaller amount of geometry to the graphics memory after a new ping and then use it as VBO and produce the large number of intersecting polygons in a geometry shader.

Illumination:

Lighting of scenes significantly enhances depth cues [9]. For this reason, we included lighting in our pipeline. A comparison for a lit and non lit visualization is shown in Figure 8. We adapted the multidirectional occlusion shading model [139] into our system for the following reasons: The model generates a high-quality soft shadowing effect in volumes without gradients. This has two advantages. First, the approximation of gradients in our pipeline would be very costly as we reconstruct the volume on the fly and voxel's neighborhood is not apriori known. Second, gradient-free lighting such as soft shadowing yield superior result over gradient-based shading in noisy data such as MRI and ultrasound [57, 139]. Furthermore, the model allows light-source specification in the hemisphere defined by the viewer and thereby allows to preset the assumed light direction which is in 12° left and $20^\circ - 30^\circ$ above the viewer [145, 109]. Finally, the model requires slice-based rendering which fits well into our pipeline. An additional example where the multidirectional occlusion shading is applied is displayed in Figure 9.

Incompatibility with a raycaster:

Even though raycasters offer more flexibility than slice-based renderers, they have several disadvantages for our setup of data compared to our approach. Even though we do not compare the quality of our results to a raycaster, we point out its implementation impracticalities which is a reason why slicing should be used. We are render the intersection geometry into the view aligned slices. Therefore, we consider our approach to be object-based. Contrarily, raycasting is considered to be image-based, because the color of each

fragment is calculated by processing the geometry. Processing the geometry for each sample along each ray is, even with optimization, infeasible compared to an object-based method. In addition, our object-based approach reconstructs slices for volume rendering. This setup allows us to apply the multidirectional occlusion shading which is one of the state-of-the-art low-frequency lighting methods.

Performance:

Table 1 lists exact performance measurements of the volume reconstruction and rendering pass of a scene composed of 128 sonar scans with resolution 434×343, such as Figure 1. We summarized measurements for shadowing and no shadowing modi, different number of slicing planes depending on the sampling distance and different scan-skipping levels k. The table indicates a significant performance boost of the implementation using geometry shaders and VBO for large geometry (128 scans, k=1). The table identifies the main bottleneck of the system – the occlusion shading. The reason is that our implementation of the model uses "ping-pong" buffers during the front-to-back slice traversal. Therefore, our future work will address optimization of this stage. Nevertheless, the listed cases indicate, that the shadow mode using a VBO achieves at least 3 fps and is able to keep up with the stream rate of the sonar (1-2 images per second). Therefore, the shadow mode can be used to view the in-situ visualization of the stream but during user interaction, fast shadow-free rendering can be used. The time which is required to update the VBO and the texture storage structure is negligible compared to the rendering time ($0-16$ms). All measurements were performed at a workstation equipped with an NVIDIA GeForce 580 GTX GPU with 1536MB graphics memory, an Intel Core i7 CPU with 3.07GHz and 12GB of RAM.

518 slicing planes				
	Shadowing		No shadowing	
	VBO:Yes	VBO:No	VBO:Yes	VBO:No
k=1	297ms	830ms	0-4ms	608ms
k=6	281ms	312ms	0-4ms	109ms
259 slicing planes				
	Shadowing		No shadowing	
	VBO:Yes	VBO:No	VBO:Yes	VBO:No
k=1	203ms	359ms	0-4ms	312ms
k=6	187ms	203ms	0-4ms	47ms

Table 1: Rendering times required for the volume-rendering and reconstruction passes measured in miliseconds (ms).

D.7 Results and discussion with domain experts

The design of our system was conducted in collaboration with a domain expert. However, in order to objectively detail the utility scenarios of our system, we conducted discussions with independent domain experts from a marine-research institution and from the hardware manufacturer who were not included in the design process. They pointed out the following main benefits of our tool.

We presented our system to Simrad, the hardware vendor of the ME70 sonar. The application was confirmed to be an *evolutionary next step in visualization of the streams of 2D sonars and the future potential for better understanding the school shape and insight in the distribution in the water column*. Scientists from Simrad found our results convincing and encouraged us to prepare our application for installation on their research vessel and to schedule with them a marine survey.

Our current results were generated from pre-recorded streams acquired during surveys of sand eel at three different locations in the North Sea in April 2010. The scans from a Simrad ME70 sonar were saved as images with associated headers and then streamed offline to our system.

Shape and spatial location of schools:

Fish species can be in some cases interpreted by the shape of their schools [119]. According to the experts, *captains currently learn to recognize fish species only using 2D projections. For example, fish schools of mackerel have elongated shape. Recognition only based on 2D projections is difficult and often fails.*

Even though it is an illegal practice, fishermen often return the catch back into the water after recognizing the misinterpretation; They must fulfill quota of fish brought to the shore stated by the law. This causes that a substantial number of fish is sacrificed. Currently, captains have to mentally reconstruct the 3D shape because there are only slicing techniques or rely on the 2D projections. Our systems will assist the learning and the interpretation process.

Early recognition of the shape allows to make decisions in real-time and adapt the survey experiments; During the survey, the scientists usually look for a school of certain species. After they have found it, they might want to take samples in order to verify the species, to define the age and size distribution of fish within the school. For both, the research and the commercial fishery, the time until the decision for the catch is made is crucial – it should be made while the vessel passes over the school or else it can be too late for a catch as the school might have moved. However, there are many factors which complicate the decision making, e.g., the sonar captures only a part of the school. An in-situ 3D visualization solution would be a great advantage.

Our tool brings clear advantages to the domain in terms of shape description. In addition, the integrated grids enable rough estimation of distances in the volume. The distance between individual grid lines corresponds to 10m in the real world. This allows to quickly judge the extent of the fish school and distance between individual density levels within the school marked by the outlines. For example, the school shown in Figure 1 spans approximately 100m. The grid is drawn at one depth level which is interactively decided by the user. Intersections with the volume allow to measure the absolute depth of

individual structures. *Such measurements help to make decision whether the school is not too small to be worth the catch or not too big so that it could destroy the vessel equipment.*

Intensity levels and temporal dimension:

The marine scientists and fishermen are interested to see different levels of density, i.e., MIP levels in the biomass: *Density distribution is important for behavioural information, species identification and necessary to know for understanding the distribution (amount and location) of the biomass value, e.g., is the main contribution only from a small part of the school or does the entire school contribute significantly.*

In Figure 1, MIP-level threshold = 0.47 and in Figure 7 = 0.44. The outlines are independent of the transfer function setup. The definition of the threshold is interactive which allows the user to browse in the MIP-levels independently of the transfer function. As the next natural extension of our system, we will include volume calculation by summing-up the thresholded samples during the rendering pass to provide the user with an exact estimate.

In fishery research, *the sonar images are always tagged with time stamps. It is essential to place the biomass in the space and in time as well. For example, when a vessel passes over the same school several times, the embedded time information resolves the ambiguity whether the fish school is one big school or a smaller school which has moved since the last observation.*

Fish swim or can be shifted by significantly by sea currents. If a current of the speed of 1mph, e.g., speed of the Gulf Stream in its slowest and widest parts in the north [106], the biomass is shifted by 26m/min which is a significant change in position. The input stream composed of images and transformations does not contain any real-time information about sea currents and fish-school movement. It is therefore impossible to compensate for these movements. Still, our system encodes the temporal dimension in the color and thickness of the outlines. The captain can relate parts of the volume to a time interval in the past and establish a degree of confidence.

With our application, we put a new powerful tool into the hands of marine scientists which has the potential to lead them to new discoveries and knowledge. In commercial fishery, it would assist with fishing-quota control and more precise decision making.

D.8 Conclusion

In this paper, we have presented a rendering system for live volume visualization of streamed 2D sonar data and their associated transformations. Our design fuses the volume reconstruction and rendering of sonar images into one step and thus reduces the number of resampling stages. The architecture of the system includes a new concept of storage with fast data access in the texture memory, geometry calculation and interpolation which is applicable for in-situ visualization of any spatially-located streamed data.

We introduced a novel concept of temporal outlines which associate parts of the volume visualization and the acquisition time of the sonar scans. Assuming that the sea environment is unstable, the temporal outlines establish a level of confidence in the volume visualization. In addition, the outlines delineate regions of maximal projected intensities higher than a user-defined threshold. Our implementation shows that this design allows

for sufficient frame rates and is thus viable for in-situ use on vessels. The prototype of the system was presented to the sonar hardware manufacturer and to the independent marine scientists. They approved it as *a natural step in the evolution in the visualization software for 2D sonars* which will certainly be beneficial for their work.

Acknowledgments

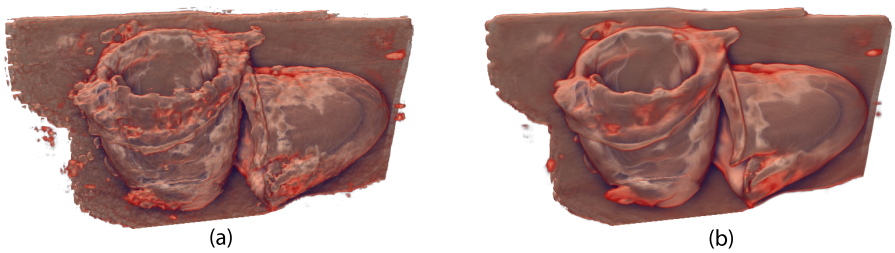
This work has been carried out within the IllustraSound research project (# 193170), which is funded by the VERDIKT program of the Norwegian Research Council with support of the MedViz network in Bergen, Norway (PK1760-5897- Project 11). Simrad Kongsberg, Horten, Norway is thanked for allowing the use of data collected by their research vessel M/S “Simrad Echo”, and financial support from The Research Council of Norway, Contract 185065; Survey Methods for Abundance Estimation of Sandeel, (SMASCC) (Ammodytes marinus) Stocks, (SMASCC) are appreciated. The authors wish to thank in particular Lars Andersen from Simrad, and the team of scientists from the Institute of marine research for their comments and fruitful discussion.

PAPER

E

Variance-Streamline Filtering for 3D Ultrasound*

Veronika Šoltészová[1], Linn Emilie Sævil Helljesen[2], Wolfgang Wein[3], Odd Helge Gilja[1,2], Ivan Viola[1,4]



(a) (b)

Figure 1: Comparison of (a) a visualization of a raw 3D ultrasound scan of a phantom dataset and (b) a visualization of the same dataset filtered with the lowest-variance streamline method.

Abstract

ULTRASOUND as an acoustic imaging modality suffers from various kinds of noise. The presence of noise especially hinders the 3D visualization of ultrasound data, both in terms of resolving the spatial occlusion of the signal by surrounding noise, and mental decoupling of signal from noise. This paper presents a novel type of structure-preserving filter that has been specifically designed to eliminate the presence of speckle and random noise from 3D ultrasound visualizations. This filter is based on a local distribution of variance for a given voxel. The lowest variance direction is assumed to be aligned with the direction of the structure. A streamline integration over the lowest-variance vector field defines the filtered output value. The new filter is compared to other popular filtering approaches and its superiority is documented on several use cases. A case study where a clinician was delineating vascular structures of the liver from 3D visualizations further demonstrate the benefits of our approach compared to the state of the art.

* This article was accepted to EG Workshop on Visual Computing for Biology and Medicine and in September 2012, it will be presented in Nörrköping, Sweden by Veronika Šoltészová. The camera ready copy of this paper might have small differences to this version.

[1] University of Bergen Bergen, Norway
[2] National Center for Ultrasound in Gastroenterology, Haukeland University Hospital, Bergen, Norway
[3] Munich University of Technology, Germany
[4] Christian Michelsen Research, Norway

E.1 Introduction

Medical ultrasound enjoys popularity as the most preferred medical modality in a number of diagnostic and medical treatment scenarios [48]. This acoustic reflectance based modality has very valuable characteristics in terms of practical bedside usage and price. Most importantly, as a live modality it has an unbeatable temporal resolution and the spatial resolution can be superior to standard 3D modalities such as computed tomography and magnetic resonance imaging. Moreover, measuring well-studied acoustic phenomena with ultrasound can provide useful information about flow, elasticity, or strain of imaged tissue. Very important for patient safety is that no ionizing radiation is associated with ultrasound examination and its usage within mechanical index limits is considered safe.

The most serious disadvantage of ultrasound imaging is the high presence of various kinds of noise that impede the image interpretation. These kinds of noise have been intensively studied and can be categorized into two distinct categories, random and structured. Structured noise can be further categorized into subcategories such as acoustic scattering (speckle), shadowing, or dropout. With regard to 2D ultrasound images, most of these noise types can be distinguished by the sonographer given substantial experience. Moreover, speckle noise is often considered as a useful source of information, and there are debates whether to keep speckle in diagnostic ultrasound imaging.

Three-dimensional ultrasound visualization is very different from the 2D case, however. In 3D renderings, random and structured noise impede the visual reconstruction of imaged structures, occlude it, modify it, and are the origin of normal perturbation that becomes a very dominant effect when calculating local illumination. Therefore, for 3D visualization the goal is to eliminate all kinds of noise and give prominence to the signal. Traditional noise removal filters, however, can potentially modify the signal up to such an extent that it is no longer diagnostically relevant. Finding a clear separation between signal and noise is not trivial and cannot be handled by common linear and non-linear filters.

The scope of the presented work is a novel structure-preserving filtering approach that is based on a local variance distribution and is designed specifically for the 3D ultrasound visualization pipeline. Figure 1 demonstrates the method on a phantom model of the myocardium [41]. Unlike traditional filters, its operator mask is a curve that locally aligns to the structure. This eliminates structure thinning or removal of structural details as is typical for other kinds of filters reviewed in Section E.2. Details on algorithmic description of the new filter are provided in Section E.3. The filtering is demonstrated on several ultrasound phantom and anatomical datasets in Section E.4 and its structure-preserving behavior is evaluated in Section E.5. Finally, the paper draws conclusions on conducted research in Section E.6.

E.2 Related work

A large body of research has been devoted to pre-processing and data enhancement for ultrasound. Sakas et al. [133] listed techniques with a good trade-off between loss of information and quality. A recent survey by Birkeland et al. [8] provides a concise overview over the ultrasound visualization pipeline, where several approaches include a

pre-filtering step prior to rendering. In this section, we review only the most relevant works related to noise reduction in ultrasound.

The following works are performing enhancement based on local data homogeneity. Kuwahara et al. [79] described a filter which divides the neighborhood of a point P into blocks and the filtered value in P was the mean value of the most homogeneous block, i.e., block with lowest variance. Karaman et al. [66] presented an adaptive filtering technique for speckle removal for B-mode ultrasound. Smoothing operators (mean or median) are applied in regions where the tissue is assumed homogeneous. These regions are obtained by region growing which is constrained only by statistical properties and the distance from the central pixel. Yanhui et al. [167] performed directional averageing based on 2D homogeneity. Pixels, which have their homogeneity above a certain threshold, remain unchanged. Other pixels are processed by their directional average filter. An edge detection is followed by directional filtering along the edge with the higher edge-value (vertical or horizontal). Farzana et al. [37] used a combination between the Euclidean distance between the origin pixel O and pixel J and a homogeneity parameter of O. This parameter is obtained from blocks of O's neighborhood which have homogeneity above a certain threshold. Bilateral filters combine pixels based on their geometric closeness and photometric similarity [149]. Viola et al. [155] presented hardware-based implementations of the median, bilateral filter, and the Kuwahara filter.

Statistical analysis of data has been facilitated for filtering and segmentation purposes in a number of works. Czerwinski et al. [28] proposed an adaptation of a median filter to solve the problem of boundary-preserving speckle reduction in ultrasound. They took a set of short lines passing through the center of a square-shaped kernel. Along each line, they computed the median. Finally, they kept the largest median value for the pixel in the center of the kernel. In their follow up work, they described how lines can be detected in ultrasound images [29]. They also discussed different methods for hypothesis testing that the actual line is going through the edge in the picture. Coupé et al. [26] adapted the nonlocal means filter [12] based on Bayesian statistics. Instead of a simple distance weighting used in the original nonlocal means filter, they used a statistical distance, i.e., Pearson Distance between two random variables X and Y that is based on their correlation $1-\rho_{XY}$.

Statistical properties were also used for determining tissue similarity in general. Patel et al. [113] used statistical moments (variance in combination with mean) for segmentation of noisy datasets. They used statistical moments to determine similarity between tissues. Their approach could be adjusted for automatic design of the operator mask. Fattal and Lischinski [38] used a variational approach for opacity classification of 3D ultrasound datasets.

Several works pursued adaptive filter design. Chinrungrueng and Suvichakorn [19] employed polynomial surface fitting to intensities. They reported results which were at least as good as results produced by a median filter but could be obtained for less computation time. Caan et al. [14] described an adaptive filtering kernel which depends on the space-variant level of noise and some similarity measures to the central pixel. The adaptation was done as a weigthed Gaussian filtering where the weights were related to the similarity of the neighborhood tensors. Eom [36] determined the shape and the orientation of the filter based on distance and angle maps, i.e., distance of a pixel to the nearest edge. The filter was then aligned to the edge tangent at the closest point of the

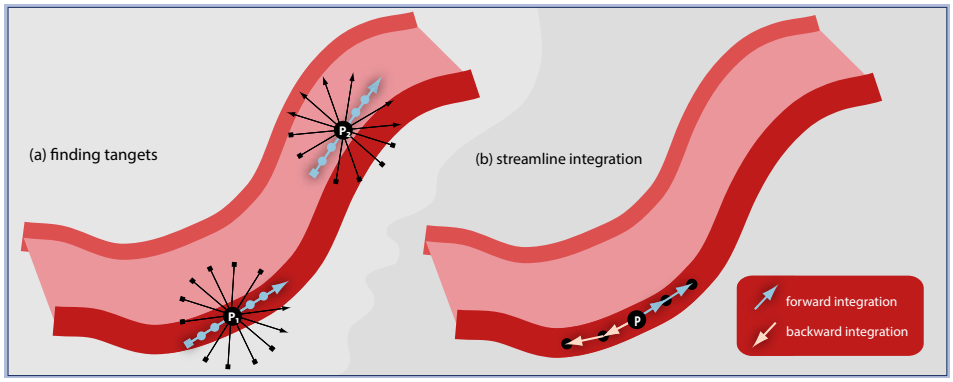


Figure 2: Determination of the line segment n=3 with the lowest variance of intensities at two points P_1 and P_2 (a). Streamline integration seeded in point P (b). Forward integration is shown in blue and backward integration in pink.

closest detected edge.

Su and Seul proposed filtering with wavelets [144]. They replaced small wavelet coefficients by zero and kept or shrank other coefficients. Anisotropic diffusion is a frequently applied filtering method which smooths inside regions but not the edges itself [114]. The edges are defined by local gradients. Michailovich and Tannenbaum [97] conducted a study where they compared performances of three nonlinear filters: wavelet denoising, total variation filtering [129] and anisotropic diffusion; and demonstrate their applicability for speckle removal in medical ultrasound.

This short review of filtering methods illustrates the large body of existing work that has addressed the problem of denoising an ultrasound signal. Still, to the best of our knowledge, there is no work that took an approach similar to our lowest-variance streamline filtering, and is specifically designed to improve 3D ultrasound visualization. It is too preliminary to claim that our technique outperforms any other form of filtering. In this paper though, we have conducted first comparative steps and relate our work to selected, most frequently used structure-preserving and denoising approaches.

E.3 Filtering Method

Speckle noise in 3D ultrasound poses a challenge to volume visualization since it obstructs interpretation and identification of structures. To improve the quality of 3D rendering, it is desired to perform a speckle-removal procedure. In data processing for the medical domain, it is very important to preserve the boundaries of structures. This is a problem of many filtering techniques: even though the boundaries and edges are preserved, they move or change shape.

There have been attempts to preserve edges by using bilateral filters using weighted averaging taking into account the distance and intensity similarity between voxels. This approach is however sensitive to noise since the intensity similarity factor is based only

on differences (local property of two points). Our approach pursues a different strategy. We are performing a selective averaging, but the selection which voxels will be taken into account is novel with respect to previous work. The filtering happens in two stages each of which can be parallelized:

1. **Determine the tangent direction** For each point P in the volume (a voxel in our context), select the direction which has the highest probability of all directions **x** to be tangent to a fictive surface going through P. The outcome of this stage is a 3D vector field.

2. **Integrate** For each point P, construct a short streamline seeded in P by integration of the vector field produced in the previous stage. The streamline defines the shape of the filtering operator mask.

With many speckle removal techniques available, our method is, to the best of our knowledge, the only technique which utilizes the principle of streamline integration to snap the filtering kernel to object boundaries. In this manner, we ensure after filtering, the boundaries have moved minimally and have preserved their intensity. In Sections E.3.1 and E.3.2, we describe each of the stages.

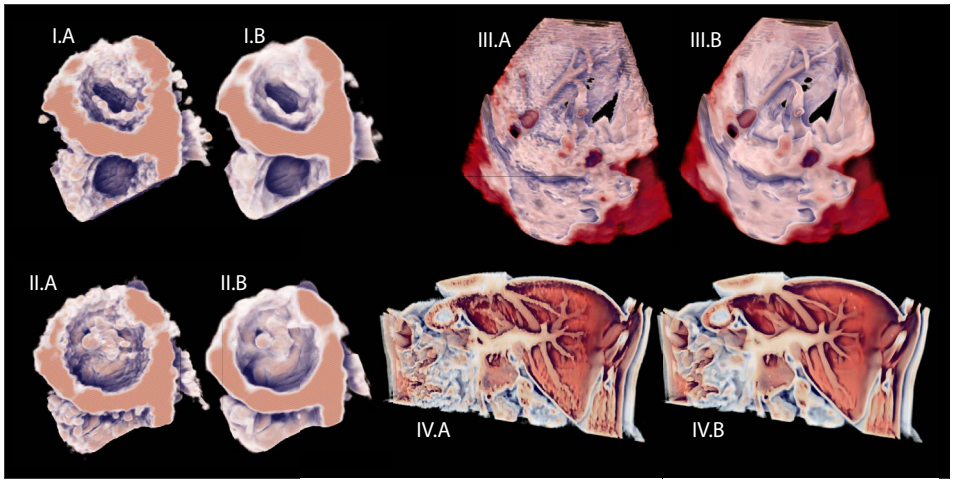


Figure 3: Comparison of non-filtered and filtered datasets from medical ultrasound and MRI: I. and II. 3D cardiac ultrasound, III. 3D ultrasound of liver, and IV. MRI of liver.

E.3.1 Local Direction of Lowest Variance

The ultrasound-inherent speckle noise poses a challenge to any local processing technique of this modality. When determining the tangent directions it is therefore necessary to evaluate a larger neighborhood, in order to find the direction of a boundary going through a voxel in noisy data. We are evaluating variance in patterns within a local neighborhood, since this is a robust measure used also in previous work for data classification [113].

We assume that values along a line segment entirely inside one tissue material will have lower variance of intensities than a line segment which is crossing several materials. To find the orientation of a line segment with lowest variance for every voxel, we proceed as follows. We align line segments centered in a point P to a discrete set of directions. These directions are obtained by rotating an initial line segment in the XY and XZ plane around P by an angle δ. This assures a minimal angular sampling rate of δ.

Each line segment is defined by the position P and the direction vector $\mathbf{x}$. Since both vectors $\mathbf{x}$ and -$\mathbf{x}$ could define the same line, we consistently select $\mathbf{x}$ with a positive y-coordinate. Then we calculate variance of the set of samples for each of the line segments. The direction $\mathbf{x}_{min}$ which corresponds to the line segment with the lowest variance will be copied to the output 3D vector field at the position of P. Formally, we define $\mathbf{x}_{min}$ as follows:

$$\mathbf{x}_{min} \mid \forall \mathbf{x}, \underset{k\in -n..n}{Var}\{f(P+k\Delta\mathbf{x})\} \geq \underset{k\in -n..n}{Var}\{f(P+k\Delta\mathbf{x}_{min})\} \tag{1}$$

where $Var\{.\}$ is the variance of a set of values, $f(P)$ is the voxel intensity at point P, Δ is a positive step size and n indicates how many samples are taken along the line segment in the positive and in the negative negative sense. In our approach, we used $n = 5$, where each $\Delta\mathbf{x}$ amounts to the size of a single voxel. The principle is shown in Figure 2a simplified in 2D. The $\mathbf{x}_{min}$ is shown in blue, all other line segments are shown in black.

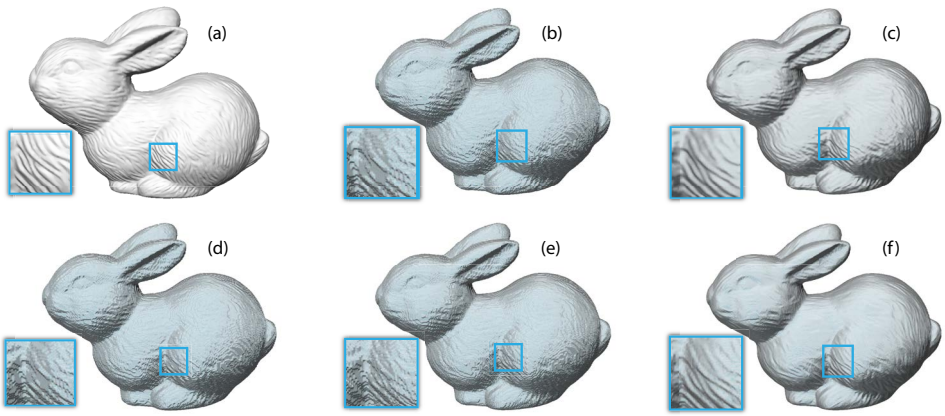


Figure 4: Mesh representation of a laser scan of a bunny (a), voxelization of the mesh (b), filtered with mean filter $3\times3\times3$ (c), median filter $3\times3\times3$ (d) Kuwahara filter $3\times3\times3$, and our method $n = 5$, $m = 5$.

E.3.2 Streamline Integration and Filtering

In the first stage, we obtained a 3D vector field where each vector represents the direction of the line segment with minimal variance. At this point, we continue with the con-

struction of the operator mask for each voxel P separately. This procedure is similar to streamline integration with P being the seed point. See also the illustration in Figure 2b.

1. **Forward integration** constructs a part of the operator mask while integrating $\mathbf{x}_{min}$ from the underlying vector field. In Figure 2b, this part of integration is marked in blue.

2. **Backward integration** uses the inverted vector field, i.e., $-\mathbf{x}_{min}$ to construct the second part of the operator mask, in Figure 2b illustrated with pink.

Both the backward- and the forward-integration parts are employing the Runge-Kutta 4 integration scheme [131, 78]. In this way, we obtain $2m+1$ samples where m is the number of integration steps. The filtered value P' at point P is then determined as the arithmetic mean of these samples:

$$P' = \frac{1}{2m+1} \sum_{i=-m}^{m} f(P_i) \tag{2}$$

With P_i being the i^{th} integration step of the streamline and $i > 0$ being the forward integration, $i = 0$ the seed point sample, and $i < 0$ backward integration. For the results presented in this paper, we used $m = 5$.

E.4 Results

We have applied our technique to various 3D ultrasound and MRI data sets. In addition, we tested the technique by filtering the staircase artifacts on voxelized meshes.

Figure 1 shows the effect of the variance-streamline filtering applied to an ultrasound scan of a phantom of a myocardium. The phantom is manufactured from a synthetic polymer called PVA (polyvinyl alcohol) which has in crystallized form acoustic properties similar to the myocardium [41]. We also pre-filtered series of ultrasound volumes capturing a human cardiac cycle. Figure 3 shows additional pairwise comparisons of filtered and not filtered datasets: 3D ultrasound of the human heart extracted from a cardiac cycle (pairs I and II), 3D liver ultrasound (III) and liver MRI (IV). The amount of speckle and noise was significantly decreased, in particular in terms of smoothed walls of the myocardium and liver vessels. The noise level in the MRI dataset decreased significantly as well while the edges remained clear.

Our filter also performs well at eliminating voxelization artifacts. We applied several frequently using filtering approaches to the voxelized mesh: mean filter, Kuwahara filter, median filter. The mean filter creates fuzzy borders. The median filter and the Kuwahara filter preserve sharp edges, but do not remove the artifacts. Our method preserves borders and at the same time, removes a large portion of the artifacts. The original mesh, as well as the voxelization and filtered datasets are compared in Figure 4.

We implemented this method as a preprocessing step in CUDA. While the second stage of the filtering process can be executed during rendering, the 3D vector field has to be precomputed as thew computation takes approximately five seconds on modern graphics hardware for a volume of size 256^3.

E.5 Case Study

In data processing and in visualization, it is highly important to maintain structures so that they represent the anatomic situation as precisely as possible. An uncareful preprocessing could cause vessel thinning or sometimes even removing of thin branches. Vessels are, however, very important because, e.g., in the liver they are influencing preoperative planning decisions and are also useful for localizing pathologies. Anatomical partitioning of the liver is determined by the liver vessel tree [25].

To assess the quality of our technique, especially its usability for filtering vessels in medical ultrasound datasets, we conducted a quantitative evaluation with a clinician specialized in gastroenterology. During the interpretation of ultrasound visualization, the clinician mentally "removes" the speckle and other kinds of noise. To compare the result of speckle removal from a visualization of non-filtered data, which is mentally filtered by a clinician, and filtered data with denoising filters, we conducted the following task.

We presented the clinician a series of visualizations of liver ultrasound without the pre-filtering. The visualizations were printed in an A4 format and put into an adhesive transparent foil. The clinician used a marker to outline the vessel tree in the liver. For each liver dataset, she received visualization of the original dataset and five pre-filtered versions using different techniques: median filter $3 \times 3 \times 3$, Kuwahara filter $3 \times 3 \times 3$, anisotropic diffusion with two distinct parameter settings, and our method. Concerning the anisotropic diffusion, it is difficult to automatically find a good parametrization (time step, κ, and the number of iterations) [42]. Therefore, we produced results with a series of combinations and selected two best settings concerning structure preserving (diffusion I) and level of noise (diffusion II). Moreover, the same parameter setting might be suitable only for the dataset it was chosen for.

Our aim was to observe, where the clinician sees the vessels in the original ultrasound dataset and compare these to the situation when this dataset is filtered with the most frequently used techniques and our technique. Figure 5 showcases the test cases including the original, i.e., the non-filtered dataset, and datasets filtered with five different techniques, including ours. Each visualization corresponds to a line drawing made by the clinician. She used a green marker to draw vessels and a red marker to draw where she was not certain about the shape of the vessel wall. For the filtered datasets the clinician was instructed to solve the task of vessel delineation only using the information extracted from the visualization. Finally, for drawing over the non-filtered dataset, she was allowed to view the 3D rendering of the dataset in an interactive application to get better insights about the structures. Therefore we consider the vascular delineation in the non-filtered dataset as the one, which is best representing the structural arrangement. For benchmarking purposes the line drawing extracted from the filtered dataset is rated according to how close it is to the line drawing from the non-filtered dataset.

In total, we evaluated three different scenarios of human vessel tree in the liver I, II, and III. All three test cases including the corresponding line drawings are listed in Figure 5. Inspired the quantitative comparison method proposed by Cole et al. [22], we evaluated the similarity between line drawings extracted from the non-filtered dataset and filtered datasets. We will refer to the drawings based on filtered data "filtered drawings" and to line drawings based on original data as to "original drawings". To obtain a simi-

Scene	Med	Kuwahara	Dif. I	Dif. II	VS
I	11.78	24.14	9.29	12.54	8.76
II	3.17	39.92	10.78	15.08	3.53
III	10.83	16.06	16.69	20.70	9.28

Table 1: Relative distance between the illustration extracted from the non filtered dataset and filtered datasets using median (Med), Kuwahara, diffusion and variance-streamline (VS) filtering.

larity measure, we first converted each line drawing to a binary image where 1 signifies line and 0 no line. Then we computed a distance field for each filtered drawing. To determine the similarity between an original drawing and a selected filtered drawing, we used element-wise multiplication of the distance field of the selected filtered drawing and the binary image of the original drawing. The sum of all values in the result image signifies the total summed error. In order to obtain a relative measure, we divided the summed error by the summed length of the corresponding filtered drawing. The summed length of a drawing is simply the number of pixels with value 1 in the binary mask. This relative measure is, in other words, an expected distance of each point on a filtered line drawing to the closest point on its corresponding original drawing. The relative distances are listed in Table 1. We can see that the variance streamline filtering has obtained the scores with the smallest relative distance to the original drawings (with one exception in scenario II).

Additionally, the clinician rated the techniques subjectively based on the following criteria:

- Are the borders clear or fuzzy? In general, it is difficult to define borders if they appear fuzzy.
- Are the borders jaggy? Finding and interpreting smooth borders is easier.
- Does the filtering mehod cause that parts of the vessel are “cut off”, but a little further, it seems that the vessel continues? Interpretation is difficult in this case.

Based on the visualizations of data, she subjectively rated the filtering methods in the following order (best to worst): variance-streamline filtering (our method), median filtering, no filtering, diffusion I, diffusion II, and Kuwahara filtering. From the evaluation we can conclude that both in our quantitative evaluation method and in subjective preference the new variance-streamline filtering method was preferred over the other techniques.

E.6 Conclusions

We described a novel filtering approach which utilized the lowest variance direction to locally identify borders of structures, and based on this information, the operator mask is locally curved. This procedure is similar to streamline integration in a vector field. We showed its applicability especially on medical ultrasound, an imaging modality that is challenging 3D visualization due to its noisiness and speckle. In addition, we evaluated how the new filtering method affects understanding of structures in ultrasound. Based on

the quantitative analysis and subjective judgment made by a clinician, we can conclude that our method preserves structures and at the same time eliminates noise which makes the interpretation of the visualization easier.

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.

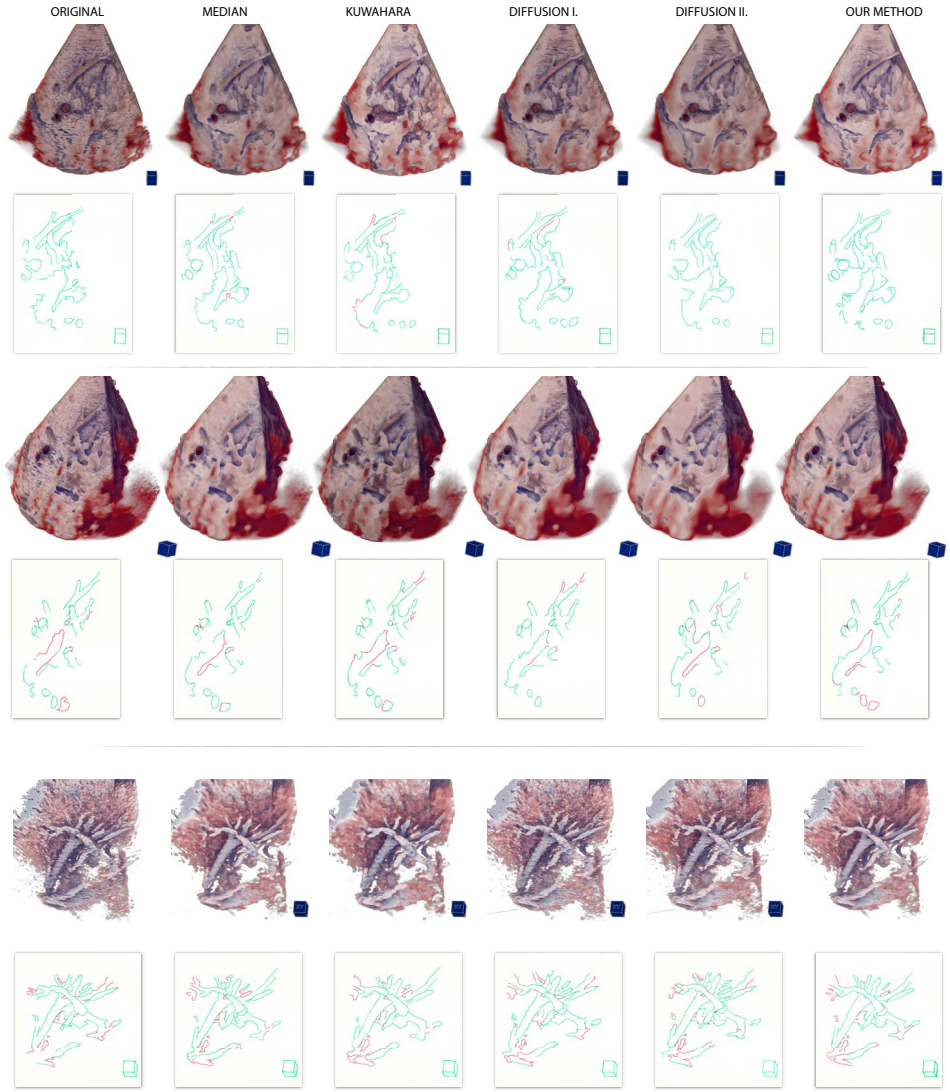
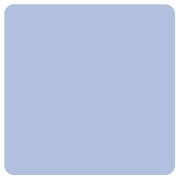


Figure 5: Comparison of the visualizations of the original dataset, median filtering $3 \times 3 \times 3$, Kuwahara filtering $3 \times 3 \times 3$, diffusion I. (Δ = 3/44, 5 iterations, κ = 12), diffusion II. (Δ = 3/44, 7 iterations, κ = 14), and our method with n=5 and m=5. Below we showcase the corresponding line drawings made by the doctor. She used a green marker except of those lines where she was rather uncertain. The cube in the bottom right was used for registration of her drawings.

Bibliography

There are many shellfish... so why not being selfish?

—Der Meister



Bibliography

[1] Adobe. Adobe Photoshop CS4 - The "Curves..." tool. `www.adobe.com/products/photoshopfamily.html`, 2008.

[2] R. C. Allen and M. L. Rubin. Chromostereopsis. *Survey of Ophthalmology*, 26(1):22–27, 1981.

[3] A. S. Bair, D. H. House, and C. Ware. Texturing of layered surfaces for optimal viewing. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1125–1132, 2006.

[4] J.-P. Balabanian, I. Viola, E. Ona, R. Patel, and E. Gröller. Sonar Explorer: A new tool for visualization of fish schools from 3D sonar data. In *Data Visualization - EuroVis 2007*, pages 155–162, 2007.

[5] U. Behrens and R. Ratering. Adding shadows to a texture-based volume renderer. In *Proceedings of IEEE Symposium on Volume Visualization 1998*, pages 39–46, 1998.

[6] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.

[7] K. Berbaum, T. Bever, and C. S. Chung. Light Source Position in the Perception of Object Shape. *Perception*, 12(5):411–416, 1983.

[8] Å. Birkeland, V. Šoltészová, D. Hönigmann, O. H. Gilja, S. Brekke, T. Ropinski, and I. Viola. The ultrasound visualization pipeline – a survey. *ArXive – e-prints*, 2012.

[9] W. Braje, G. Legge, and D. Kersten. Invariant recognition of natural objects in the presence of shadows. *Perception*, 29(4):383–398, 2000.

[10] S. Bruckner and M. E. Gröller. VolumeShop: An interactive system for direct volume illustration. In *Proceedings of IEEE Visualization*, pages 671–678, 2005.

[11] S. Bruckner and M. E. Gröller. Enhancing Depth-Perception with Flexible Volumetric Halos. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1344–1351, 2007.

[12] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 60–65, 2005.

[13] M. Bunnel. *GPU-Gems*, volume 2, chapter Dynamic Ambient Occlusion and Indirect Lighting. Addison Wesley, 2005.

[14] M. Caan, G. Khedoe, D. Poot, A. den Dekker, S. Olabarriaga, , K. Grimbergen, L. van Vliet, and F. Vos. Adaptive noise filtering for accurate and precise diffusion estimation in fiber crossings. In *Proceedings of Medical Image Computing and Computer–Assisted Intervention*, pages 167–174, 2010.

[15] W. Cai and G. Sakas. Data intermixing and multi-volume rendering. *Computer Graphics Forum*, 18(3):359–368, 1999.

[16] P. Cézanne. Le panier de pommes. Art Institute, Chicago, 1894.

[17] P. Cézanne. Nature morte avec rideu et pichet fleuri. The Hermitage Museum, St. Petersburg, 1899.

[18] Y. Chen, R. Yin, P. Flynn, and S. Broschat. Aggressive region growing for speckle reduction in ultrasound images. *Pattern Recognition Letters*, 24(4-5):677–691, 2003.

[19] C. Chinrungrueng and A. Suvichakorn. Fast edge-preserving noise reduction for ultrasound images. In *IEEE Nuclear Science Symposium Conference Record*, volume 3, pages 18/99–18/103, 2000.

[20] J. Cohen. *Statistical power analysis for the behavioral sciences*. Routledge Academic Press, New York, 2^{nd} edition, 1988.

[21] J. Cohen. A power primer. *Psychological Bulletin*, 112(1):155–159, 1992.

[22] F. Cole, A. Golovinskiy, H. S. B. Limpaecher, A. Finkelstein, T. Funkhouser, and S. Rusinkiewicz. Where do people draw lines? *ACM Transactions on Graphics*, 27(3), 2008.

[23] 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.

[24] G. Coleman. Landscapes. `http://garycolemanartist.com/landscapes{_}main.htm`, 2011.

[25] C. Couinaud. *Le foie. Études anatomiques et chirurgicales*. Masson et Cie, Paris, 1957.

[26] P. Coupé, P. Hellier, C. Kervrann, and C. Barillot. Nonlocal means-based speckle filtering for ultrasound images. *IEEE Transactions on Image Processing*, 52(5):2221–2229, 2009.

[27] P. Coupé, P. Hellier, X. Morandi, and C. Barillot. Probe Trajectory Interpolation for 3D Reconstruction of Freehand Ultrasound. *Medical Image Analysis*, 2007.

[28] R. Czerwinski, D. Jones, and W. O'Brien Jr. Ultrasound speckle reduction by directional median filtering. In *Proceedings of the International Conference on Image Processing*, volume 1, pages 358–361, 1995.

[29] R. Czerwinski, D. Jones, and W. O'Brien Jr. Line and boundary detection in speckle images. *IEEE Transactions on Image Processing*, 7(12):1700–1714, 1998.

[30] L. da Vinci. Virgin of the Rocks. Louvre, Paris, 1486.

[31] L. da Vinci. Virgin of the Rocks. National Gallery, London, 1508.

[32] C. Dachsbacher, M. Stamminger, G. Drettakis, and F. Durand. Implicit visibility and antiradiance for interactive global illumination. *ACM Transactions on Graphics*, 26(3):61.1–61.10, 2007.

[33] E. De Haan, R. Erens, and A. Noest. Shape from shaded random surfaces. *Vision Research*, 35(21):2985–3001, 1995.

[34] P. Desgranges, K. Engel, and G. Paladini. Gradient-free shading: A new method for realistic interactive volume rendering. In *Proceedings of Vision, Modeling, and Visualization*, pages 209–216, 2005.

[35] W. Einthoven. Stereoscopie durch Farbendifferenz. *Von Graefes Archiv*, 19:211–238, 1885.

[36] K. B. Eom. Speckle reduction in ultrasound images using nonisotropic adaptive filtering. *Ultrasound in Medicine and Biology*, 37(10):1677–1688, 2011.

[37] E. Farzana, M. Tanzid, K. Mohsin, M. Bhuiyan, and S. Hossain. Adaptive bilateral filtering for despeckling of medical ultrasound images. In *Proceedings of TENCON - IEEE Region 10 Conference*, pages 1728–1733, 2010.

[38] R. Fattal and D. Lischinski. Variational classification for visualization of 3D ultrasound data. In *Proceedings of IEEE Visualization 2001*, pages 403–410, 2001.

[39] J. Faubert. Seeing depth in colour: More than just what meets the eyes. *Vision Research*, 34(9):1165–1186, 1994.

[40] R. Fleming, A. Torralba, and E. Adelson. Specular reflections and the perception of shape. *Journal of Vision*, 4:798–820, 2004.

[41] H. Fosså, P. Lunde, and K. Matre. Ultrasound phantom for myocardium, 2011.

[42] L. Fritz. Diffusion-based applications for interactive medical image segmentation. In *Proceedings of the Central European Seminar on Computer Graphics (CESCG)*, 2006.

[43] A. Gallardo. Lambertian shading. In *3D Lighting: History, Concepts and Techniques*, page 117. Charles River Media, Inc., Massachusetts, 2001.

[44] GE Healthcare. Voluson E8, HDlive imaging – experience realism with volume ultrasound, 2012. www.gehealthcare.com/euen/ultrasound/products/general-imaging/voluson-e8/hdlive/index.html.

[45] GE Healthcare. A closer look at GE's pocket-sized Vscan ultrasound, Sep 13th, 2010. www.gereports.com/a-closer-look-at-ges-pocket-sized-vscan-ultrasound.

[46] S. Geisser and S. W. Greenshouse. On methods in the analysis of profile data. *Psychometrika*, 24:95–112, 1959.

[47] J. J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.

[48] O. H. Gilja, T. Hausken, S. Ødegaard, Ø. Wendelbo, and M. Thierley. Mobile ultrasonography in a medical department. *Tidskrift for Norsk Lægeforening*, 19:270–285, 2003.

[49] E. Gombrich. *Art and Illusion*. Princeton University Press, 1969.

[50] A. Gooch, B. Gooch, P. Shirley, and E. Cohen. A non-photorealistic lighting model for automatic technical illustration. *Proceedings of ACM SIGGRAPH*, pages 447–452, 1998.

[51] B. Gooch, E. Reinhard, and A. Gooch. Human facial illustration: Creation and psychophysical evaluation. *ACM Transactions on Graphics*, 23(1):17–44, 2004.

[52] B. Gooch, P.-P. J. Sloan, A. Gooch, P. Shirley, and R. Riesenfeld. Interactive technical illustration. In *Proceedings of the 1999 symposium on Interactive 3D graphics*, pages 31–38, 1999.

[53] T. Grossman and R. Balakrishnan. An evaluation of depth perception on volumetric displays. *AVI '06:Proceedings of the working conference on Advanced visual interfaces*, page 193, 2006.

[54] M. Hadwiger, A. Kratz, C. Sigg, and K. Bühler. GPU-accelerated Deep Shadow maps for Direct Volume Rendering. In *Proceedings of SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware*, pages 49–52, 2006.

[55] E. Hecht. *Optics*. Addison Wesley, 2^{nd} edition, 1987.

[56] F. Hernell, P. Ljung, and A. Ynnerman. 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*, pages 105–112, 2008.

[57] F. Hernell, P. Ljung, and A. Ynnerman. Local Ambient Occlusion in Direct Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics*, 99(2), 2009.

[58] F. Hernell, A. Ynnerman, and P. Ljung. Efficient Ambient and Emissive Tissue Illumination using Local Occlusion in Multiresolution Volume Rendering. In *Proceedings of Volume Graphics 2007*, pages 1–8, 2007.

[59] D. Hönigmann, J. Ruisz, and C. Haider. Adaptive design of a global opacity transfer function for direct volume rendering of ultrasound data. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, pages 489–496, 2003.

[60] 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.

[61] J. M. Jacobs, R. Hammerman-Rozenberg, Y. Maarvi, A. Cohen, and J. Stessman. The impact of visual impairment on health, function and mortality. *Aging Clinical and Experimental Research*, 17(4):281–286, 2005.

[62] A. Johnston and P. Passmore. Shape from shading. I: Surface curvature and orientation. *Perception*, 23:169–189, 1994.

[63] D. Jönsson, E. Sundén, A. Ynnerman, and T. Ropinski. State of the art report on interactive volume rendering with volumetric illumination. In *Eurographics STAR program*, pages 53–74, 2012.

[64] A. Joshi, J. Caban, P. Rheingans, and L. Sparling. Case study on visualizing hurricanes using illustration-inspired techniques. *IEEE Transactions on Visualization and Computer Graphics*, 15:709–718, 2009.

[65] A. Karamalis, W. Wein, O. Kutter, and N. Navab. Fast hybrid freehand ultrasound volume reconstruction. In *Proceedings of SPIE Medical Imaging*, pages 726114.1–726114.8, 2009.

[66] M. Karaman, M. Kutay, and G. Bozdagi. An adaptive speckle suppression filter for medical ultrasonic imaging. *IEEE Transactions on Medical Imaging*, 14(2):283–292, 1995.

[67] D. Kersten, P. Mamassian, and D. Knill. Moving cast shadows and the perception of relative depth. Technical Report 6, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany, 1994.

[68] Y. Kim and A. Varshney. Persuading visual attention through geometry. *IEEE Transactions on Visualization and Computer Graphics*, 14(4):772–782, 2008.

[69] M. Knecht. State of the art report on ambient occlusion. Technical report, Technische Universität Wien, Vienna, Austria, 2007.

[70] J. Kniss, G. Kindlmann, and C. Hansen. Multidimensional transfer functions for interactive volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 8(3):270–285, 2002.

[71] J. Kniss, S. Premoze, C. Hansen, P. Shirley, and A. McPherson. A Model for Volume Lighting and Modeling. *IEEE Transactions on Visualization and Computer Graphics*, 9(2):150–162, 2003.

[72] J. Koenderink, A. van Doorn, and A. Kappers. Surface perception in pictures. *Perception and Psychophysics*, 52(5):487–496, 1992.

[73] J. Koenderink, A. van Doorn, and A. Kappers. Ambiguity and the *mental eye* in pictorial relief. *Perception*, 30:431–448, 2001.

[74] R. J. Korneliussen, Y. Heggelund, I. K. Eliassen, O. K. Øye, T. Knutsen, and J. Dalen. Combining multibeam-sonar and multifrequency-echosounder data: examples of the analysis and imaging of large euphausiid schools. *ICES Journal of Marine Science: Journal du Conseil*, 66(6):991–997, 2009.

[75] K. Krissian, C. F. Westin, R. Kikinis, and K. G. Vosburgh. Oriented speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 16(5):1412–1424, 2007.

[76] J. Krivanek, J. Ferwerda, and K. Bala. Effects of Global Illumination Approximations on Material Appearance. *Proceedings of ACM SIGGRAPH*, 29(4), 2010.

[77] T. Kroes, F. H. Post, and C. P. Botha. Exposure render: An interactive photorealistic volume rendering framework. *PLoS ONE*, 7(7):e38586, 2012.

[78] M. W. Kutta. Beitrag zur näherungsweisen Integration totaler Differentialgleichungen. *Zeitschrift für Mathematik und Physik*, 46:435–453, 1901.

[79] M. Kuwahara, K. Hachimura, S. Eiho, and M. Kinoshita. Processing of riangiocardiographic images. In *Digital processing of biomedical images*, pages 187–202. Plenum Press, New York, USA, 1976.

[80] M. Levoy. Display of Surfaces from Volume Data. *IEEE Computer Graphics and Applications*, 8:29–37, 1987.

[81] F. Lindemann and T. Ropinski. Advanced Light Material Interaction for Direct Volume Rendering. pages 101–108, 2010.

[82] 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.

[83] M. Livingstone. *Vision and art – the biology of seeing*. Abrams, paperback edition, 2008.

[84] T. Luft, C. Colditz, and O. Deussen. Image enhancement by unsharp masking the depth buffer. In *Proceedings of ACM SIGGRAPH*, pages 1206–1213, 2006.

[85] D. N. MacLennan and E. J. Simmonds. *Fisheries Acoustic*. Chapman Hall, London, 1992.

[86] P. Mamassian and R. Goutcher. Prior knowledge on the illumination position. *Cognition*, 81:B1–9, 2001.

[87] P. Mamassian and D. Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351–2367, 1996.

[88] É. Manet. La rue mossier aux paveurs. http://www.wikipaintings.org/en/edouard-manet/the-road-menders-rue-de-berne-1878, 1878.

[89] D. Margulis. *Photoshop LAB Color: The Canyon Conundrum and Other Adventures in the Most Powerful Color space*. Peachpit Press, 2005.

[90] MathWorks. Matlab: The language of technical computing. www.mathworks.com, 2012.

[91] H. Matisse. Autoportrait. Statens Museum for Kunst, Copenhagen, 1906.

[92] J. W. Mauchly. Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11(2):204 –209, 1940.

[93] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.

[94] L. Mayer, Y. Li, and G. Melvin. 3D visualization for pelagic fisheries research and assessment. *ICES Journal of Marine Science: Journal du Conseil*, 59(1):216–225, 2002.

[95] S. McCloud. Chapter 4 - time frames. In *Understanding Comics - The Invisible Art*, pages 94–117. HarperCollins Publishers, 1993.

[96] À. Méndez-Feliu and M. Sbert. From obscurances to ambient occlusion: A survey. *The Visual Computer*, 25(2):181–196, 2009.

[97] O. Michailovich and A. Tannenbaum. Despeckling of medical ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(1):64–78, 2006.

[98] E. Mingolla and J. Todd. Perception of solid shape from shading. *Biological Cybernetics*, 53:137–151, 1986.

[99] O. A. Misund. Dynamics of moving masses: variability in packing density, shape, and size among herring, sprat, and saithe schools. *ICES Journal of Marine Science: Journal du Conseil*, 50(2):145–160, 1993.

[100] C. Monet. La pie. http://en.wikipedia.org/wiki/File:Monet_-_The_Magpie.jpg, 1869.

[101] C. Monet. Les meules. http://en.wikipedia.org/wiki/Haystacks_(Monet), 1901.

[102] C. Monet. Série des cathédrales de Rouen. http://en.wikipedia.org/wiki/Rouen_Cathedral_(Monet), 1902-1904.

[103] T. R. Nelson and T. T. Elvins. Visualization of 3D ultrasound data. *IEEE Computer Graphics and Applications*, 13(6):50–57, 1993.

[104] T. K. Nguyen, A. Eklund, H. Ohlsson, F. Hernell, P. Ljung, C. Forsell, M. T. Andersson, H. Knutsson, and A. Ynnerman. Concurrent volume visualization of real-time fMRI. In *IEEE/EG Volume Graphics*, pages 53–60, 2010.

[105] 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.

[106] NOS. National Ocean Service: Ocean facts - how fast is the Gulf Stream? `http://oceanservice.noaa.gov/facts/gulfstreamspeed.html`, March 2011.

[107] S. Ødegaard, O. H. Gilja, and H. Gregersen. *Basic and New Aspects of Gastrointestinal Ultrasonography*. Advanced Series in Biomechanics. World Scientific, 2005.

[108] S. Ødegaard, L. B. Nesje, and O. H. Gilja. *Atlas of Endoscopic Ultrasonography*. Fagbokforlaget, 2007.

[109] 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.

[110] J. Paramo, F. Gerlotto, and C. Oyarzun. Three dimensional structure and morphology of pelagic fish schools. *Journal of Applied Ichthyology*, 26(6):853–860, 2010.

[111] J. M. Parramón. *The book of color*. Watson-Guptil Piblications, New York, english edition, 1993.

[112] D. Patel, S. Bruckner, I. Viola, and M. E. Gröller. Seismic volume visualization for horizon extraction. In *Proceedings of IEEE Pacific Visualization*, pages 73–80, 2010.

[113] D. Patel, M. Haidacher, J.-P. Balabanian, and M. E. Gröller. Moment curves. In *Proceedings of the IEEE Pacific Visualization Symposium 2009*, pages 201–208, 2009.

[114] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. In *Proceedings of IEEE Computer Society Workshop on Computer Vision*, pages 16–22, 1987.

[115] J. D. Pfautz. Depth perception in computer graphics. *Technical Report*, 2002.

[116] M. Pharr and G. Humphreys. *Physically Based Rendering – from theory to implementation*. Morgan Kaufmann, second edition, 2010.

[117] B. T. Phong. Illumination for computer generated pictures. *ACM Commununications*, 18(6):311–317, 1975.

[118] 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.

[119] T. J. Pitcher and J. K. Parrish. Functions of shoaling behavior in teleosts, in behavior of teleost fishes. In *Behavior of Teleost Fishes*, pages 363–440. Chapman Hall, London, 1996.

[120] Z. Pizlo and M. Salach-Golyska. 3-D shape perception. *Perception and Psychophysics*, 57(5):695–714, 1995.

[121] R. Prager, A. Gee, and L. Berman. Stradx: real-time acquisition and visualization of freehand three-dimensional ultrasound. *Medical Image Analysis*, 3(2):129–140, 1999.

[122] R. Prager, A. Gee, G. Treece, and L. Berman. Freehand 3D ultrasound without voxels: volume measurement and visualisation using the Stradx system. *Ultrasonics*, 40(1-8):109–15, 2002.

[123] M. Puerta. The power of shadows: shadow stereopsis. *Journal of the Optical Society of America. A, Optics and image science*, 6(2):309–311, 1989.

[124] V. S. Ramachandran. Perception of shape from shading. *Nature*, 331:163–166, 1998.

[125] C. Rezk-Salama. GPU-Based Monte-Carlo Volume Raycasting. In *Proceedings of Pacific Graphics 2007*, pages 411–414, 2007.

[126] T. Ritschel. Fast GPU-based visibility computation for natural illumination of volume data sets. In *Short Papers Eurographics 2007*, pages 17–20, 2007.

[127] T. Ropinski, C. Döring, and C. Rezk-Salama. Interactive Volumetric Lighting Simulating Scattering and Shadowing. In *Proceedings of IEEE Pacific Visualization*, pages 169–176, 2010.

[128] T. Ropinski, J. Meyer-Spradow, S. Diepenbrock, J. Mensmann, and K. H. Hinrichs. Interactive Volume Rendering with Dynamic Ambient Occlusion and Color Bleeding. *Computer Graphics Forum*, 27(2):567–576, 2008.

[129] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[130] M. Ruiz, I. Boada, M. Feixas, and M. Sbert. Viewpoint information channel for illustrative volume rendering. *Computers & Graphics*, 34(4):351–360, 2010.

[131] C. Runge. Über die numerische Auflösung von Differentialgleichungen. *Mathematische Annale*, 46:167–178, 1895.

[132] S. Rusinkiewicz, M. Burns, and D. DeCarlo. Exaggerated shading for depicting shape and detail. *ACM Transactions on Graphics*, 25(3):1199–1205, 2006.

[133] G. Sakas, L. Schreyer, and M. Grimm. Preprocessing and volume rendering of 3D ultrasonic data. *IEEE Computer Graphics and Applications*, 15(4):47–54, 1995.

[134] M. Schott, V. Pegoraro, C. Hansen, K. Boulanger, J. Stratton, and K. Bouatouch. A Directional Occlusion Shading Model for Interactive Direct Volume Rendering. *Computer Graphics Forum*, 28(3):855–862, 2009.

[135] P. M. Shankar. Speckle reduction in ultrasonic images through a maximum likelihood based adaptive filter. *Physics in Medicine and Biology*, 51(21):5591–5602, 2006.

[136] P. Shanmugam and O. Arikan. Hardware accelerated ambient occlusion techniques on GPUs. In *Proceedings of Symposium on Interactive 3D Graphics and Games*, pages 73–80, 2007.

[137] Simrad. ME70. `www.simrad.com`, March 2011.

[138] P.-P. J. Sloan, J. Kautz, and J. Snyder. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM Transactions on Graphics*, 22(3):376–381, 200.

[139] V. Soltészová, D. Patel, S. Bruckner, and I. Viola. A multidirectional occlusion shading model for direct volume rendering. *Computer Graphics Forum*, 29(3):883–891, 2010.

[140] 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.

[141] E. Steen and B. Olstad. Volume rendering of 3D medical ultrasound data using direct feature mapping. *IEEE Transactions on Medical Imaging*, 13(3):517–25, 1994.

[142] A. J. Stewart. Vicinity Shading for Enhanced Perception of Volumetric Data. In *Proceedings of IEEE Visualization*, pages 355–362, 2003.

[143] A. Stompel, E. B. Lum, and K.-L. Ma. Visualization of multidimensional, multivariate volume data using hardware-accelerated non-photorealistic rendering techniques. In *Proceedings of Pacific Graphics*, pages 394–402, 2002.

[144] C. Su and H. Seul. A speckle reduction filter using wavelet-based methods for medical imaging application. In *Proceedings of the* 23^{rd} *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 3, pages 2480–2483, 2001.

[145] J. Sun and P. Perona. Where is the sun? *Nature Neuroscience*, 1:183–184, 1998.

[146] E. Sundén, A. Ynnerman, and T. Ropinski. Image plane sweep volume illumination. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2125–2134, 2011.

[147] J. Todd. The visual perception of 3D shape. *Trends in Cognitive Science*, 8(3):115–121, 2004.

[148] 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.

[149] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the 6th conference on Computer Vision*, pages 839–846, 1998.

[150] E. R. Tufte. The visual display of quantitative information. In *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[151] A. J. van Doorn and J. Koenderink. The influence of environmental cues on pictorial relief. *Perception - ECVP Abstract Supplement*, 29, 2000.

[152] V. van Gogh. Prisoners exercising. `http://en.wikipedia.org/wiki/File:Vincent_Willem_van_Gogh_037.jpg`, 1890.

[153] 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.

[154] R. Vergne, R. Pacanowski, P. Barla, X. Granier, and C. Shlick. Improving shape depiction under arbitrary rendering. *IEEE Transactions on Visualization and Computer Graphics*, 17(8):1071–1081, 2011.

[155] I. Viola, A. Kanitsar, and M. E. Gröller. Hardware-based nonlinear filtering and segmentation using high-level shading languages. In *Proceedings of IEEE Visualization*, pages 309–316, 2003.

[156] L. C. Wanger, J. A. Ferwerda, and D. P. Greenberg. Perceiving Spatial Relationships in Computer-Generated Images. *IEEE Computer Graphics and Applications*, 12(3):44–58, 1992.

[157] C. Ware. More on contours. In *Information Visualization - Perception for Design*, pages 198–205. Morgan Kaufmann, 2004.

[158] C. Ware and G. Franck. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics*, 15(2):121–140, 1996.

[159] C. Weigle and D. Banks. A Comparison of the Perceptual Benefits of Linear Perspective and Physically-Based Illumination for Display of Dense 3D Streamtubes. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1723 –1730, 2008.

[160] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.

[161] A. Wilkie, R. F. Tobler, C. Ulbricht, G. Zotti, and W. Purgathofer. An Analytical Model for Skylight Polarisation. In *Proceedings of the Eurographics Symposium on Rendering*, pages 387–399, 2004.

[162] G. Windholz. The second signal system as conceived by Pavlov and his disciples. *Pavlovian Journal of Biological Science*, 25(4):163–173, 1990.

[163] H. Winnemöller, D. Feng, B. Gooch, and S. Suzuki. Using NPR to evaluate perceptual shape cues in dynamic environments. In *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, pages 85–92, New York, NY, USA, 2007. ACM.

[164] J. Woodring, C. Wang, and H.-W. Shen. High dimensional direct rendering of time-varying volumetric data. In *Proceedings of IEEE Visualization*, pages 417–424, 2003.

[165] YafaRay. Yafaray 0.0.9: Yet another free raycaster. `www.yafaray.org`, 2008.

[166] R. Yagel, A. Kaufman, and Q. Zhang. Realistic volume imaging. In *Proceedings of IEEE Visualization*, pages 226–231, 1991.

[167] G. Yanhui, H. Cheng, T. Jiawei, and Z. Yingtao. A novel approach to speckle reduction in ultrasound imaging. *Ultrasound in Medicine and Biology*, 35(4):628–640, 2009.

[168] I. Yu, A. Cox, M. H. Kim, T. Ritschel, T. Grosch, C. Dachsbacher, and J. Kautz. Perceptual influence of approximate visibility in indirect illumination. *ACM Transactions on Applied Perception*, 6:24:1–24:14, 2009.

[169] S. Zhukov, A. Iones, and G. Kronin. An Ambient Light Illumination Model. In *Rendering Techniques*, pages 45–56, 1998.