FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Human Visual Perception of 3D Surfaces

A Gauge-Figure-Based Surface Perception Study

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Pascal Plank
Matrikelnummer 1225804

an der Fakultät für Informatik
der Technischen Universität Wien

Betreuung: Assistant Prof. Dipl.-Ing. Dr. techn. Ivan Viola
Mitwirkung: Dipl.-Ing. Dr. techn. Matthias Bernhard
Dipl.-Ing. Dr. techn. Nicholas Waldin

Wien, 21. Juni 2015

Pascal Plank	Ivan Viola

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.ac.at

FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Human Visual Perception of 3D Surfaces

A Gauge-Figure-Based Surface Perception Study

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Pascal Plank
Registration Number 1225804

to the Faculty of Informatics
at the Vienna University of Technology

Advisor: Assistant Prof. Dipl.-Ing. Dr. techn. Ivan Viola
Assistance: Dipl.-Ing. Dr. techn. Matthias Bernhard
Dipl.-Ing. Dr. techn. Nicholas Waldin

Vienna, 21st June, 2015

Pascal Plank	Ivan Viola

Erklärung zur Verfassung der Arbeit

Pascal Plank
Hanreitergasse 3/3/37, 1210 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 21. Juni 2015

Pascal Plank

Acknowledgements

Special thanks to Matthias Bernhard and Ivan Viola for the excellent collaboration on this project as well as the their constructive feedback and suggestions for this thesis.

This work was supported through grants from the Vienna Science and Technology Fund (WWTF) through project VRG11-010 and by the EC Marie Curie Career Integration Grant through project PCIG13-GA-2013-618680.

Kurzfassung

Der Gauge-Figure-Task ist ein methodisches Werkzeug, um die Wahrnehmung der Oberflächenorientierung durch einen Betrachter in Renderings und Visualisierungstechniken zu studieren. Ursprünglich entwickelt um die Wahrnehmung von Bildern, d. h. nicht wahrheitsgetreue Stimuli, zu untersuchen, wurde der Gauge-Figure-Task seither genutzt, um absolute Wahrnehmungsfehler durch den Vergleich von Oberflächenschätzungen eines Beobachters mit der Ground-Truth der Flächennormalen zu messen.

In dieser Bachelorarbeit wurde die absolute Genauigkeit des Gauge-Figure-Tasks untersucht, d. h. wie gut die Oberflächenschätzungen mit den tatsächlichen Flächennormalen übereinstimmen. Um den Messfehler zu isolieren wurde eine User Study unter Verwendung verschiedener primitiver Objekte durchgeführt. Weiters wurden mehrere Depth Cues bereitgestellt, einschließlich Depth-from-Motion und Stereodisparität, um potentielle Wahrnehmungsfehler zu minimieren. Es wurde erwartet, dass eine stereoskopische Darstellung der Gauge-Figure den Wahrnehmungsfehler ebendieser drastisch reduzieren würde. Während der Experimente sammelte ich über 16.300 Messungen von 17 Teilnehmern unter verschiedenen Betrachtungsbedingungen, bei denen entweder die Stimuli, die Gauge-Figure, beide oder keines in stereo dargestellt wurden.

Die Ergebnisse zeigen, dass die Gauge-Figure-Schätzungen für primitive Stimuli, z. B. eine Kugel oder ein Zylinder, bei Modalität-konsistenten Betrachtungsbedingungen gut mit der Ground-Truth übereinstimmen, d. h. wenn Stimuli und Gauge-Figure beide in stereo oder beide in mono dargestellt wurden. Im Gegensatz dazu wurde eine enorme Slant-Unterschätzung festgestellt wenn die Gauge-Figure in stereo und die Stimuli in mono dargestellt wurden. Zusätzlich gilt im umgekehrten Fall, d. h. Gauge-Figure in mono und die Stimuli in stereo, dass selbst für einfache Objekte der Slant überschätzt wird.

Diese Bachelorarbeit umfasst allgemeine Hintergrundinformationen und frühere Arbeiten zu diesem Thema, das Design, den Aufbau und die Vorgehensweise bei der User Study, sowie die Ergebnisse und eine qualitative Beurteilung dieser. Weiters werden zwei alternative Erklärungen für die gefundenen Ergebnisse diskutiert und ein Ausblick auf mögliche zukünftige Arbeiten gegeben.

Abstract

The gauge figure task is a methodological tool to study an observer's perception of surface orientations in renderings and visualization techniques. Originally developed to probe the perception of paintings, i. e. not veridical stimuli, the gauge figure task has since been used to measure absolute perceptual errors by comparing an observer's surface estimates with the ground truth surface normals.

In this bachelor thesis the accuracy of the gauge figure task was investigated, i. e. how well the probed surface estimates align with the perceived surface normals on an absolute scale. To isolate the probing error a user study was carried out using different primitive objects and several depth cues, including depth-from-motion and stereo disparities, to minimize potential perceptual errors. It was expected that a stereoscopic presentation of the gauge figure would reduce the perceptual error of the gauge figure dramatically. During the experiments I collected about 16.300 probes from 17 participants under different viewing conditions where either the stimuli, the gauge figure, both or none of them were presented in stereo.

The results show that the gauge figure estimates for primitive stimuli, e. g. a sphere or a cylinder, align well with the ground truth in modality-consistent conditions, i. e. where stimuli and gauge figure were both presented in stereo or both in mono. In contrast to this, a gauge figure presented in stereo to probe monoscopic stimuli resulted in an enormous slant underestimation. In addition, in the inverse case, where the gauge figure is presented in mono and the stimuli in stereo, an overestimation occurred - even for simple stimulus objects.

This bachelor thesis covers the general background and previous work for this subject, the design, setup and procedure of the user study as well as the results and a qualitative assessment. Furthermore, two alternative explanations for the found results are discussed and an outlook for possible future work is given.

Contents

Kurzfassung ix

Abstract xi

Contents xiii

1 Introduction 1

2 Background 3
- 2.1 Stereopsis . . . 4
- 2.2 Perception of surfaces and shape . . . 6

3 Previous Work 9
- 3.1 Psychophysical measuring techniques . . . 9
- 3.2 Previous results of psychophysical experiments . . . 12

4 Surface perception user study 19
- 4.1 Stimuli . . . 19
- 4.2 Experiment Setup and Viewing Conditions . . . 23
- 4.3 Sampling points and participants . . . 25
- 4.4 User study procedure . . . 26

5 Results and Discussion 29
- 5.1 Results . . . 29
- 5.2 Qualitative Assessment . . . 31
- 5.3 Discussion . . . 34

6 Conclusion and Future Work 37

7 Appendix A: ANOVA 39

Bibliography 41

Glossary 45

CHAPTER 1

Introduction

Computer Graphics and Visualizations take advantage of the human visual system to form mental visions in order to make something visible to the mind. But how can you ensure that everybody forms the same mental image when one wants to convey something visually, e. g. the surface of an object? This discrepancy of real stimuli versus perceived stimuli is the general topic of this thesis.

In 1992 Koenderink et al. proposed a methodological tool called the "gauge figure task" [KvDK92]. Originally intended to study the perception of shapes on photographs, the gauge figure task has become the most commonly used method to measure a user's perception of surface normals in virtual 3D scenes. In computer graphics surface normals are used to determine a surface's orientation towards a light source for shading. Since then several studies revealed that surface normals are usually perceived incorrectly by human observers[DHEN95] [MK96] [STPV12].

Some methods have been developed to improve human visual perception, e.g. Pineo and Ware proposed a computational model of human vision using neural network simulations to automatically evaluate and optimize visualizations [PW12]. Solteszova et al. proposed a statistical shading model to improve normals on Lambertian shaded surfaces for better shape perception [STPV12]. Both approaches improved perception, but some inconsistencies between 3D object and perceived object remained.

So far it is unclear how much of the deviation from the ground truth is caused by the gauge figure task itself and in extension whether it is reliable enough as a measuring tool. The goal of this bachelor thesis is to give an overview of human visual perception, tools to measure it and to quantify the error and accuracy limitations caused by the gauge figure method. For this, a gauge figure user study was performed. Apart from this bachelor thesis the results of the conducted user study were also published as a coauthored technical report along with Matthias Bernhard, Manuel Waldner, Veronika Solteszova and Ivan Viola [BWP+15].

CHAPTER 2

Background

Perception is the process by which humans form a mental representation of their environment and interpret the world around them. There are many theories of human perception, but most define it as recognizing, organizing and interpreting sensory information of any kind. In general it deals with human senses such as sight, hearing, touch, smell, taste, thermoception, nociception (sense of physiological pain), balance, etc. Humans can by far perceive the most information through sight. Researchers at the University of Pennsylvania School of Medicine estimated that the human retina can transmit visual input at about 8.75 megabits per second, which is roughly the same rate as an ethernet connection [KMS$^+$06]. In this thesis the focus lies on visual perception.

The human visual perception can be divided into low-level and high-level aspects. High-level aspects include the principles of gestalt laws, e. g. the law of proximity, which states that objects that are close together are perceived as a group. Low-Level aspects on the other hand are simple visual properties such as shape, movement, spatial position and color. Those simple visual cues are being preattentively processed by the brain. Preattentive processing is intrinsic and uncontrollable compared to the attentive or controlled perceptual processing. It is fast and processes visual stimuli in parallel, usually within 250 ms. Preattentively perceived visual cues such as color are processed by sensory memory, often called iconic memory, which allows for the fast detection of the mentioned visual attributes. This fast way of processing visual information is somewhat evolutionary hardwired into the lower levels of the human visual system [WGK10].

The traditional model of human mental function has been that the senses provide separate data to the brain, which are translated into the appropriate mental phenomena, e. g. auditory experience into a melody, visual images into a mountains, etc. Contrary to this "bottom-up" process many studies suggest that a second "top-down" process occurs simultaneously, meaning that the sensual perceptions are also shaped by the conceptual understanding of the world and the current context [Koe29] [WGK10]. Without this understanding of the world in a preexisting network of distinct and interrelated ideas and concepts, the sensory experience of the world would be undifferentiated chaos rather than the discrete objects we actually apprehend. Danko Nikolic introduced the term "ideasthesia" to describe this parallel process [Nik09].

As a result of the bottom-up and top-down processes human perception is context-dependent and subjective. This has many practical implications, e. g. in the design of visualization systems [STPV12] as well as the perception of shapes and surfaces of 3D objects like in the conducted experiments for this bachelor thesis. One of the goals of the experiments for this thesis was to maximize the top-down influence in order to improve perception as much as possible to identify the errors in perception introduced exclusively by the measuring method itself. This is for example achieved by using a proper lighting setup or providing animations of the presented objects [DHEN95] [MK96]. Further details of the experiment setup are described in Chapter 4.

2.1 Stereopsis

An important aspect of human vision is stereopsis or stereo vision, which describes the perception of depth and the 3D impression of objects. Humans have, like many other animals, two eyes located at different lateral positions, usually between 52 and 78 mm apart [GBC$^+$88]. In each eye slightly different images are projected onto the retina. These positional differences, mainly in the relative horizontal position of objects, are referred to as "binocular disparities". The disparities are processed by the visual cortex in the brain, which allows for the extraction of depth information of objects in a given scene [HHR95].

When the two fields of view - one from each eye - overlap, there are two ways the brain can deal with this confusion: the two images can be fused ("binocular fusion") or one image can be suppressed so that only the other one is seen. The phenomenon that an object is seen doubly in two images is called "double vision" or "diplopia", Figure 2.1 shows an exemplary image of diplopia. The fusion of images occurs only in a small volume of visual space around a fixated point, outside of this "Panum's fusional area" double vision occurs. The condition that a person cannot see in 3D using stereo vision is called "stereoblindness". Affected individuals see as if they are only using one eye or "monocular vision" [HHR95].

Figure 2.1: Example of diplopia. (Image courtesy Wikimedia Commons)

Normal computer monitors only display one image at a time, which is seen by both eyes simultaneously. This does not agree with the natural way humans perceive their environment visually. To provide slightly separate images for each eye several techniques have been developed, e. g. anaglyph 3D, polarized 3D systems and active shutter 3D systems. The basic concept of all of these examples is the same: each eye has to receive a slightly different image simulating the natural seeing process.

In active shutter 3D systems images are presented for the left eye while blocking the right eye's view and then the right-eye image is presented while blocking the left eye's view. By repeating this process quickly it creates two slightly different and smooth images, one for each eye.

The active shutter technology was used for this user study. I utilized the stereoscopic gaming kit from Nvidia with LC shutter glasses (liquid crystal shutter glasses). In LC shutter glasses each glass contains liquid crystals, which become opaque when voltage is applied. If no voltage is applied the crystals are transparent. The glasses are synchronized with the refresh rate of the screen by an infrared timing signal. Screens for shutter glasses have to be able to work at twice the normal refresh rate (120 instead of 60 Hertz) since they have to display twice as many images for a given time interval compared to normal monitors.

There are two main factors to recreate a proper 3D impression: convergence and parallax. Convergence describes the inward movement of both eyes towards each other to maintain single binocular vision when viewing an object [CSR84]. This movement is usually expressed as an angle, the higher the angle value is, the nearer the observed object is to the eyes, Figure 2.2. When the convergence is fixed, any object between the eyes and the convergence point is closer, while objects beyond the convergence point are farther away from an observer. A convergence higher than about 6 degrees feels uneasy, i. e. the object is too close. If the value is too small, the stereo sensation will be lost because the object is too far away [Cra].

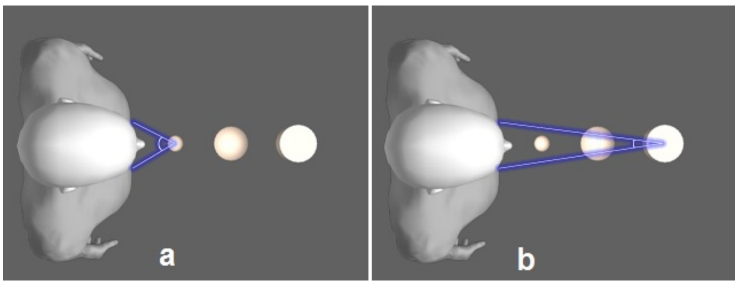

Figure 2.2: Image "a" shows high convergence (the object is near) and image "b" shows low convergence (object is distant). [Cra]

Parallax images, in contrast to motion parallax, describe the two separate images passing through to the left and right eye. This creates the illusion of depth in any 3D media. There are three types of parallax: positive, zero and negative. Positive parallax happens when the object offsets to the right in the right image and to the left in the left

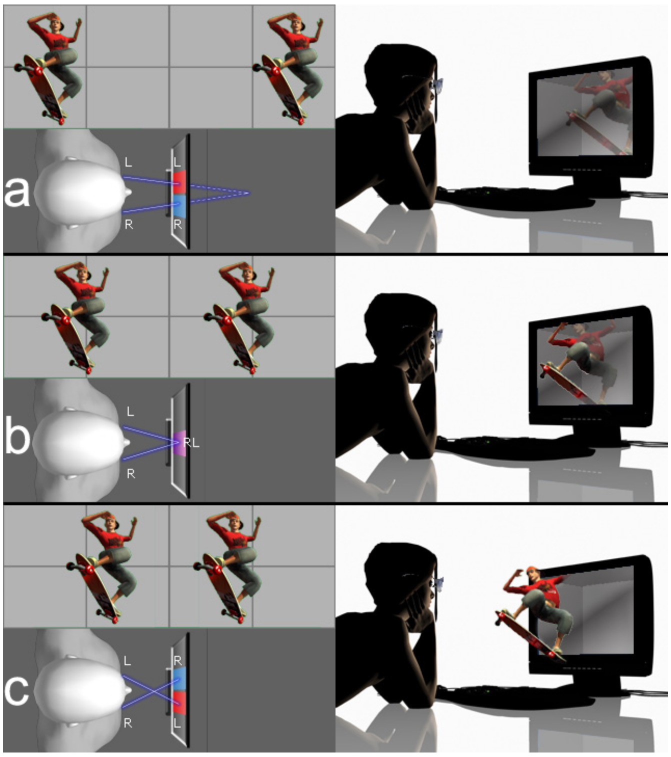

Figure 2.3: The top two images (a) show positive parallax. The next two images (b) show zero parallax, and the two bottom images (c) illustrate negative parallax where the object appears to come out of the screen. [Cra]

image. The binocular focus falls behind the display. Negative parallax is the opposite of positive parallax and occurs when an object offsets to the left in the right image and vice versa. The binocular focus falls in front of the display and the object appears to ”come out of the monitor”. Zero parallax happens when the parallax images are superimposed on the display, the binocular focus falls on the same display, which is called the plane of zero parallax. Zero parallax was used in the conducted user study, since it reduces possible distortions of the presented objects, while conveying a good 3D impression [Sch13]. Figure 2.3 shows the three types of parallax.

2.2 Perception of surfaces and shape

The process of surface perception, more precisely the perception of surfaces of 3D objects displayed on a 2D screen, contains several stages and is visualized in Figure 2.4. The starting point is a 3D object represented as a mesh of triangles in a virtual scene. Through transformations the scene is projected onto an image plane and rendered. The result of this process is a picture displayed on a 2D monitor in ”screen space”.

The screen emits light, which is partly gathered by the eye-lenses and projected onto the retina. On the retina some preattentive processes take place called retinal processing, which eventually generate signals that are sent to the brain via the optic nerve. The signals from both eyes converge in the primary visual cortex in the brain where higher visual processes take place. As mentioned in Chapter 2, seeing is a dual process involving not only the incoming signals from both eyes, but also preexisting concepts and ideas about our world. The result of these bottom-up and top-down processes is represented in "mental space" or "pictorial space" [KvDKT01].

The term "space" might wrongly indicates a real 3D space. Actually there is no real space. Pictorial space rather refers to the three-dimensional spatial impression inside our heads. It consists of the visual field (the 2D image created on the retina) and the perceived depth created by the brain. The shape of surfaces in "pictorial space" is usually referred to as "pictorial relief" [KvDKT01].

This is commonly seen as the last step in the perception pipeline, but by using probing techniques such as the gauge figure task to extract properties of a surface in pictorial space, e.g. a surface normal, we can create something called gauge space. With estimated surface normals in gauge space it is possible to study how well the mental image in pictorial space aligns with the actual surface of the visualized virtual object [CSD+09], i.e. how accurate a surface is perceived.

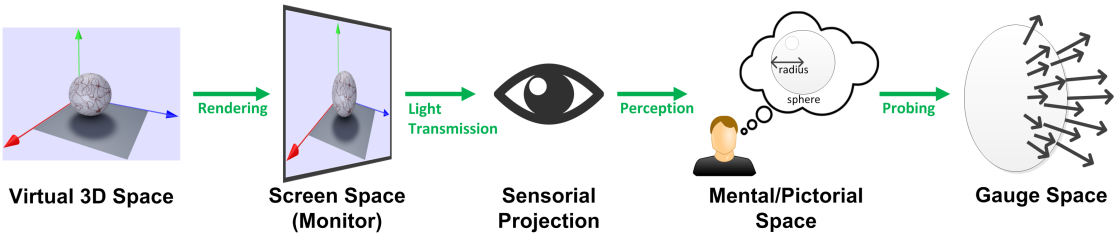

Figure 2.4: Perception Pipeline from a virtual 3D object to gauge space.

Zimmerman et al. concluded that the mental component of perception involves at least two representations of space [ZLC94]. First, a local representation of surface orientation derived from pictorial cues, i.e. pictorial space, and second, a global representation of observer-centered distance derived from binocular disparity and motion parallax.

Belhumeur et al. describe an important visual aspect called the "bas-relief ambiguity" [BKY97]. There is an implicit ambiguity in determining an unknown object's real 3D structure, when it is viewed orthographically and with Lambertian reflectance. The shading as well as the shadows on an object with surface $f(x, y)$ are identical to the shading and shadowing on any generalized bas-relief transformation of this object, $f'(x, y) = \lambda f(x, y) + \mu x + \nu y$. For truly identical shading the object's albedo has to be transformed as well. Furthermore, even small movements of the object or of the viewer cannot resolve this ambiguity in determining the flattening or scaling in λ. Notice that this is only true for orthographical projection of Lambertian shaded objects, if specular highlights are introduced or perspective projection is used this effect is not present anymore. Figure 2.5 illustrates this effect. Despite the bas-relief ambiguity and other visual ambiguities, the human visual system is able to counteract many of these perceptual inaccuracies, leading to a somewhat accurate impression of the world around us. Chapter 5.1 describes in more detail how accurate and precise the participants of the conducted user study actually were and where their perception was significantly wrong in comparison to the ground truth.

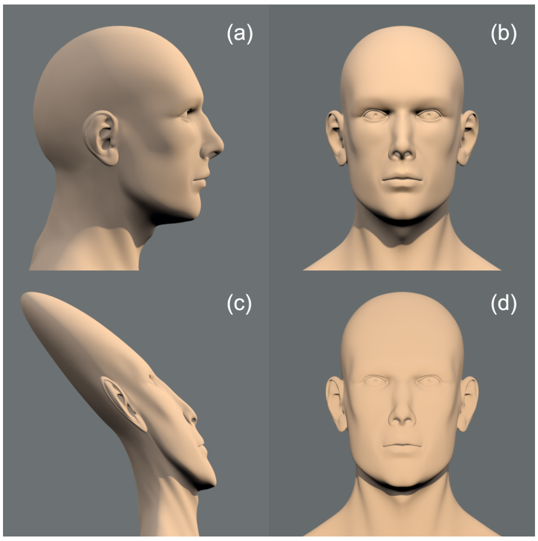

Figure 2.5: Example of bas-relief ambiguity with a Lambertian shaded face under orthographical projection. (a) and (b) show an undistorted face from right and front, (c) and (d) show the same face after strong shearing and scaling in z direction. Notice that (b) and (d) appear to have the same geometry. Only the changed lighting conditions and albedo in the image may hint that the face was transformed, but the general impression of face geometry remains unchanged.

CHAPTER 3

Previous Work

3.1 Psychophysical measuring techniques

First experiments in probing human perception of 3D objects can be traced back to the 19th century [Tod04]. These early experiments investigated stereoscopic vision by using small points of light presented in total darkness. Among other discoveries the experiments revealed that an observers' perception can be distorted in a way, so that physically straight lines appear to be curved. Since then psychophysical measuring techniques have become more refined, but the main task remains the same. In general the techniques have in common that observers have to estimate some aspect of local 3D structure at many different probing points on an object's surface to measure what an individual actually perceives. Nowadays there are many techniques for measuring a user's perception of surfaces. Koenderink et al. and Todd summarized the three most frequently used methods [KvDKT01] [Tod04], which are also explained in more detail in the following paragraphs.

3.1.1 Relative depth probe task

This method uses shaded surfaces as stimuli to measure depth perception [Tod04]. Two points on the surface are marked with differently colored dots. The observer has to choose which point is perceived closer in depth by pressing a key that corresponds to a color. Solteszova et al. used this method to evaluate the quality for an improved shadowing technique in visualizations [SPV11]. Figure 3.1 shows an example stimulus for this type of probing task.

3.1.2 Depth-profile adjustment mask

On each trial of this task a shaded object is shown with a linear arrangement of colored dots as an overlay [Tod04]. The same arrangement of dots is also presented on a second

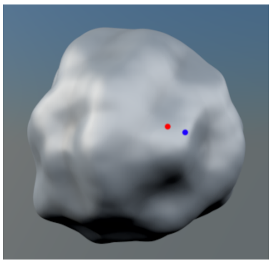

Figure 3.1: Example stimulus for the relative depth probe task with the blue dot being nearer to the viewer than the red one.

display with a blank background. On the second display each of the dots can be moved by mouse interaction to align the arrangement with the perceived surface height profile in the first image. Figure 3.2 shows an example for this probing method.

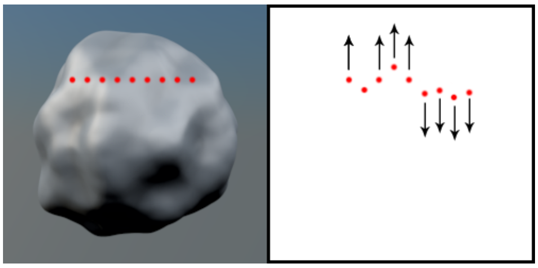

Figure 3.2: Example of the depth-profile adjustment mask method.

3.1.3 The Gauge Figure Task

The gauge figure method by Koenderink et al. has shown to be an intuitive, fast and relatively easy method to probe a 3D object's perceived surface normals [KvDK92]. In this method, the observers' task is to adjust a gauge figure in such a way that it appears to lie flat on the object's surface in pictorial space. The gauge figure is an orthographically projected Tissot's indicatrix, i. e. an ellipse indicating the amount of distortion, with a stick perpendicular to the plane defined by the indicatrix. When the gauge figure is aligned with a surface that is perpendicular to the viewing direction, it is seen as a circle. If the surface is slanted from the viewing direction, it appears as an ellipse. For a perfectly adjusted indicatrix the stick is aligned with the surface normal at the point where it intersects the surface as shown in Figure 3.3. The length of the stick is equal to the radius of the circle.

To quantify a surface normal of the local tangent plane, i. e. the depth gradient ∇Z, in pictorial space the normal's slant and tilt are used. Slant describes the angle between the surface normal and the view vector, and tilt the azimuth direction of the surface normal in the eye space. In other words slant describes the orientation in depth, and tilt

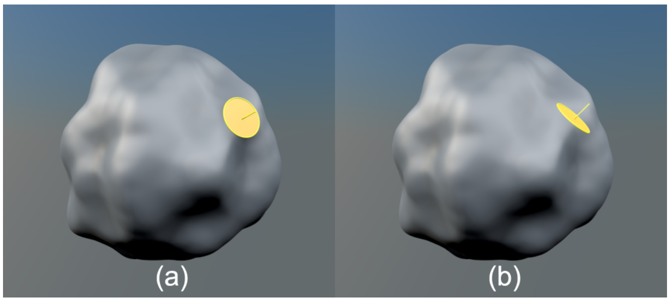

Figure 3.3: (a) a perfectly aligned gauge figure, (b) a badly adjusted gauge figure.

describes the orientation of the slant in the image plane (compass direction). Slant σ can go from 0 to $\pi/2$ or 90 degrees, and tilt τ can go from 0 to 2π or 360 degrees.

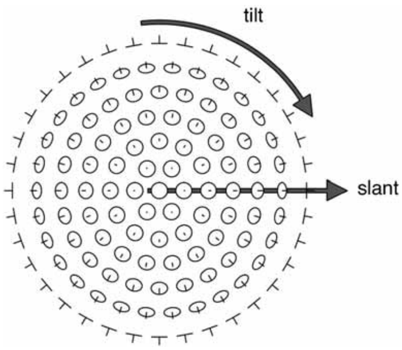

Figure 3.4: Different slants and tilts of a normal. Contrary to the discrete illustration, slant and tilt are both continuous. [vK04]

The possible directions form a hemisphere with its pole at the viewing axis as shown in Figure 3.4. The pole has a slant $\sigma = 0$ and an arbitrary tilt τ [Ste83a] [Ste83b], since the angular deviation from the viewing axis is 0 and any rotation of a view-axis aligned normal around this axis results in an identical normal. If the tip of the indicatrix-stick is adjusted to meet the indicatrix-circle, the gauge figure represents a slant of 45 degrees, since the length of the stick and the radius of the disc are equal [KvDK92].

The design of the original gauge figure task experiment consists of four stages as shown in Figure 3.5. First, one has to select a suitable image for the experiment, the experimenter then has to determine the outline of the object represented in the image, e. g. manually or by using edge detection algorithms. Afterwards the measurement samples are defined by a triangle grid overlay within the object's contour. For this a triangulated 2D plane covering the entire screen is constructed, but only triangles entirely within the outer contour of the object are used for the final experiment. The denser the grid, the more samples are going to be estimated. In the actual gauge figure task the grid is invisible and only serves as a source for sample point coordinates x_i and y_i.

At this point the participants are invited to perform the gauge figure task at the defined measurement points to obtain their surface normal estimates. Usually the observers' heads rest on chin rests in order to assure equal distance from the screen and to position the eyes centered in regards to the display. An equal configuration for all observers allows comparisons among the participants' results.

The last step is the 3D reconstruction using the observers' estimates. In this process the triangles of the 2D grid are adjusted in 3D to fit the estimated surface normals [Wij12]. As a result the grid, which is usually textured with the estimated object's image, represents the user estimated and therefore perceived 3D structure of the object.

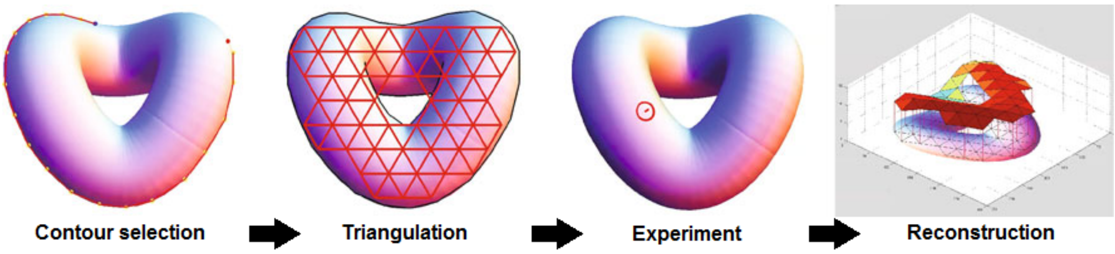

Figure 3.5: The four stages of the original gauge figure task design. [Wij12]

The original gauge figure task was designed to probe the surface perception of images and paintings, but with virtually constructed 3D objects, it is possible to compare the user estimated surface normals to the ground truth.

In comparison to the other two measuring techniques the gauge figure task is by far the easiest and the most natural. The judgment involves no overt reasoning and is therefore immediate, because participants do not have to form image abstractions from perception. In comparison to the other probing methods, the gauge figure task is the most reliable [KvDKT01] [Wij12].

3.2 Previous results of psychophysical experiments

Since experiments in probing human perception date way back into the 19th century, the accumulated body of knowledge about this topic is vast and some studies only differ in minuscule details yet came to new and sometimes different conclusions than their predecessors. In this chapter an overview over the most relevant previous findings in the area of human visual perception of surfaces and shapes is given. These findings are also partly used in the user study experiments to improve participants' perception as much as possible.

3.2.1 Slant-Underestimation and influence of Lambertian shading

Zimmerman et al. performed experiments to measure the accuracy of human perception of surface slants using 3D test planes, which were projected perspectively onto a 2D screen [ZLC94]. They calculated the perceived slant for four observers from estimates of the relative length of two orthogonal lines on a surface. The results showed that the **slant was on average underestimated by 7 degrees** with an average standard deviation of 3.35 degrees, but no systematic differences in the standard deviations were found. They also found that depth estimates of disconnected surfaces are usually not accurate, and that perception involves at least two representations of space, one of which is called "pictorial space" nowadays.

In another series of experiments Mamassian and Kersten found an **underestimation of the perceived slant by as much as 30 degrees when the displayed slant was 60 degrees** [MK96]. To be more precise **observers tended to underestimate the perceived slant for slants larger than 20 degrees** and **an overestimate under this threshold**. In contrast to the findings of Koenderink et al. [KvDK92], they also found a larger variance for the perceived tilt than for the perceived slant. Furthermore, participants were less biased at estimating the surface orientations when the surface shape was locally egg shaped rather than having a saddle or cylindrical shape.

They also concluded that the direction of illumination is presumably not used by the participants to estimate the shape of the object, which agrees with Mingolla and Todd [MT86]. Another interesting finding of Mamassian and Kersten was that it seems that if Lambertian shading was used in the experiment rather than only silhouettes, the observers were more accurate with less bias or less variability. Even though their participants complained that the silhouettes did not appear 3D, the performance was not worse than with fully shaded object. The authors also concluded that the **slant underestimation could be the result of depth overestimation** or a flattening of the textured objects. A second explanation is that the participants might have misperceived the global orientation of the objects.

De Haan et al. investigated how much of the surface ambiguity is left unconstrained by pure shading information [DHEN95]. For this they used images of random shading patterns on elliptical Gaussian hills and valleys illuminated by a single light source and ambient light. Observers had to adjust 3D local probes and the projections of these probes were then superimposed onto the shown images. The experiments revealed that despite of the large theoretical ambiguity in the stimuli the settings were reproducible and showed considerable agreement among observers. By comparing the estimates to the real surface normals they were able to identify several perceptual biases, the strongest of which are that the **global surface slant is systematically underestimated and that shading alone cannot convey the depth of a scene correctly**.

Solteszova et al. confirmed the previous findings of systematic slant underestimation and developed a perceptual-statistics shading model for Lambertian shaded surfaces [STPV12]. This model should compensate the incorrect perception of surface normals in order to improve visualizations. For this they used a dataset of about 275.000 solved gauge figure trials accomplished by a total of 560 people, which was published by Cole et al. [CSD+09]. The publication includes user responses, scene settings, documentation and the dataset itself with results for both fully-shaded and for line drawing conditions. Solteszova et al. only used the estimates for fully shaded surfaces.

From analyzing this dataset they could determine that the **crossing point of over- and underestimation of surface slants lies between 15 to 25 degrees**, confirming the findings of Mamassian and Kersten with a crossing point at about 20 degrees [MK96]. Further analysis of the dataset revealed that slants with a tilt going north tend to be underestimated less than average, for slants with south-tilts the opposite is true. The slant of normals pointing east and west are perceived very closely to the average underestimation. According to Solteszova et al. these findings are consistent with the conclusion of Todd that the underestimation of slants cannot be compensated by scaling in depth but rather by a shearing transformation in depth [Tod04]. For the new shading model a 2D map of the slants and tilts was generated. By inverting this map a look-up map was generated to adjust the surface normals during the shading process, Figure 3.6.

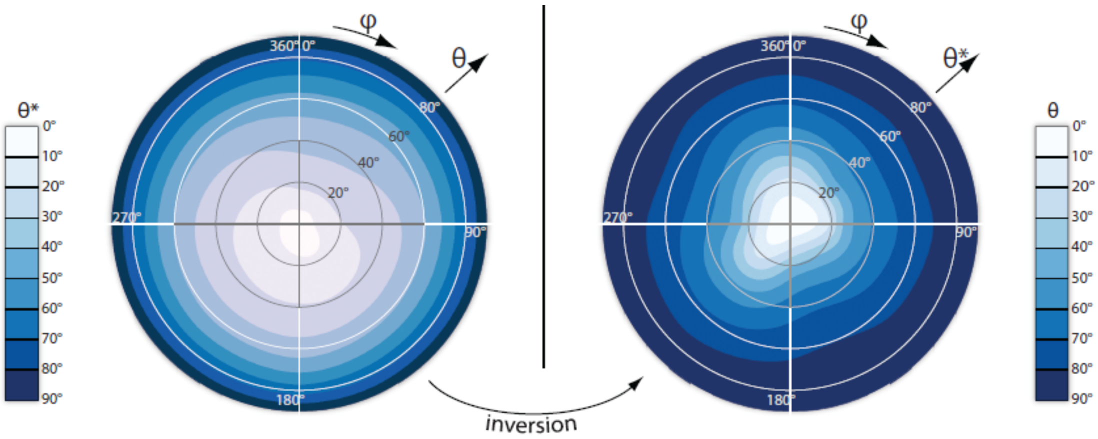

Figure 3.6: On the left the estimated slants θ and tilts φ are shown in comparison to the ground truth normals as a color-coded map. The right image is the inverse of the left and is used as a look-up map for the adjustment of the surface normals. Notice that in the left image the slants with north facing tilts are better estimated than the slants with south facing tilts. [STPV12]

To assess the efficiency of their new model, they repeated an evaluation experiment with 40 participants and confirmed that the error in perception decreased, mainly in the [40, 60] degree slant interval where the distorted perception is highest [STPV12]. Since the authors modified shading, they also modified the luminance in the final appearance of objects, which in turn affects depth perception [TI12]. The manipulation of surface normals also changes the appearance of materials, making objects appear shinier.

3.2.2 Influence of texture, brightness and age

Johnston and Passmore found that the presence of texture increases curvature discrimination thresholds [JP94]. Threshold in this case is defined as the standard deviation of the error distribution. This problem of recovering object shape from shaded and textured objects may resulted from difficulties in separating changes in brightness levels due to geometry from those due to mere surface markings. They also concluded that the surface curvature is calculated directly from depth cues present in the retinal images rather than in higher cognitive processes. Furthermore, **curvature discrimination is best in the case of shaded objects displayed on stereoscopic screens**, and the introduction as well as the enhancement of **binocular disparity cues improves slant discrimination**.

As two experiments conducted by Norman and Wiesemann revealed, the general **ability to perceive surface normals from shading and specular highlights** seems to be relatively robust and **continues to function effectively at least till the age of 80** [NW07]. In their experiments the authors made use of the gauge figure task on randomly generated and smoothly curved 3D objects and used participants from two age groups with mean ages of about 24 and 71 respectively. In the first experiment binocular disparity (stereo vision) with shutter glasses was used and in the second experiment objects were rotated in depth. Significant, but small effects of age were only found in experiment one. The observers were also able to **judge surface tilts more precisely than the slant component**, which confirms the findings of Mamassian and Kersten [MK96], but also contradicts the results of Koenderink et al. [KvDK92]. **Stereo vision also seems to improve the estimates** of both the older and the younger participants with and without texture applied to the objects, confirming the findings of Johnston and Passmore [JP94].

Koenderink et al. also measured pictorial relief for smooth solid objects using the gauge figure task [KvDC96]. The scene was always identical, but the renderings varied in illumination. This led to systematic alterations in pictorial relief. Additionally, **brighter parts in the stimuli were interpreted by observers as being nearer in pictorial space than darker areas**.

3.2.3 Presumed light position and glossy surfaces

To resolve convex-concave ambiguities in the perception of surface curvatures, the human visual system has to guess the position of the main light source. In a series of experiments Sun and Perona found that the **preferred light direction is from above and slightly shifted to the left in regards to the viewer** [SP96]. This assumed light position is also used by artists in their paintings throughout history.

Liu and Todd investigated several possible perceptual biases under a variety of rendering conditions such as shadows, specular highlights and surface inter-reflections [LT04]. The conducted experiments revealed that there seems to be a **strong perceptual bias towards perceiving images as convex** rather than concave, and the **preferred illumination position is from above**, confirming the results of Sun and Perona [SP96]. Furthermore, shadows and/or specular highlights did not seem to improve the accuracy of observers' estimates of the sign of a curved surface above mere chance, but when those aspects of normal shading were included, the performance improved significantly. Also, convex surfaces produced a greater perceived depth than concave ones with similar relief.

O'Shea et al. conducted a gauge figure experiment using smooth Lambertian shaded 3D shapes without shadows under monocular vision [OBA08]. The 3D shape was a "blobby sphere" viewed through a physical aperture. They used a physical aperture rather than a virtual one to guarantee that the participants would not mistake the edges of the aperture for the silhouette of the object. They found that the participants' **estimated surface normals were more accurate and precise when the light was positioned above the viewpoint**. At an angle of 20-30 degrees between light and viewing direction, the errors were minimized, confirming previous results that the assumed light position is above and slightly to the left in regards to the observer [SP96] [LT04].

Another interesting result in conjunction with perception of surfaces is that the direction of light seems to affect the perception of glossy and matte surfaces. Faisman and Langer found that if one increases the slant of the light source to twice that of the surface slant angle, the participants perception of shape improves for glossy surfaces [FL13]. Additionally, **for high slant angles glossy surfaces were better estimated than matte ones**. Todd and Mignolla also found that **shiny surfaces enhance the perception of curvature**, although this does not seem to have any effect on the perceived direction of illumination [TM83].

3.2.4 Monoscopic vs. stereoscopic vision

In regards to the perception of objects under stereoscopic vision Buelthoff and Mallot found that even **in the absence of edges, disparate shading creates a more vivid stereoscopic depth perception** [BM88]. If the disparities are completely removed, i.e. shape from shading, the perceived depth is reduced significantly, which was already concluded by Haan et al. [DHEN95]. If edge information is available, it seems to override both the information from shape from shading and disparate shading.

Besides "inventing" the gauge figure task Koenderink et al. also conducted experiments using a synopter [KvDK94]. A synopter is a type of binocular viewer by which both eyes receive exactly the same image. Figure 3.7 shows a sketch of a synopter. For most observers the perceived depth range increases dramatically, if one uses the device to look at a flat image. To some observers the effect is even stronger than that of regular stereoscopy. When real scenes and not flat scenes are viewed with the synopter they take on a flat appearance like if ones closes an eye. The problem with flat monitor screens is that the visual cues such as occlusion or perspective projections are 3D, but the resulting image is presented at a 2D screen. Therefore the **brain receives somewhat conflicting information: 3D cues from the rendering and 2D cues from the eyes looking at a flat 2D monitor**. This conflict is called the "depth-cue conflict". In their work Koenderink et al. concluded that **human vision uses a compromise** between the flatness of the picture surface and the relief due to monocular cues. In synoptic viewing the monocular relief outweighs the flatness of the picture, whereas in binocular viewing the flat picture surface outweighs the relief.

Vishwanath and Hibbard examined the widely held belief that stereopsis is a by-product of binocular vision [VH13]. They found empirically that this is in fact not the case, but the qualitative characteristics associated with stereopsis can occur for static 2D pictures without binocular vision too. The authors concluded that stereopsis is a

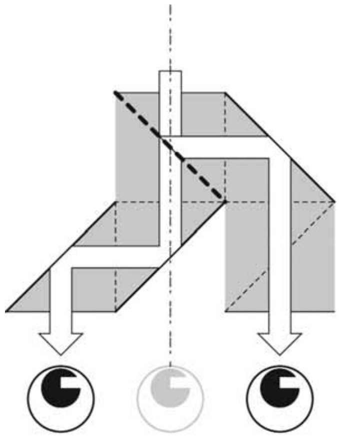

Figure 3.7: A sketch of the synopter design. The incoming light is split by a semi-transparent mirror so that both eyes receive the same light information. Notice that the length of both light rays are the same inside the synopter. [vK04]

measurable qualitative attribute. The induction stereopsis while viewing pictures is according to the authors not consistent with explanations based on depth-cue conflicts or the perception of greater depth magnitudes. In their experiments they found that a significant majority of participants had a **better impression of depth under monocular viewing than under binocular viewing**, which confirms the findings of Koenderink et al. [KvDK94].

Wardle et al. conducted a series of experiments examining the influence of edges in stereo vision [WPG14]. They found that edges are important for proper stereoscopic slant perception and that both the depth information from monocular geometry and binocular disparity might interact to resolve 3D scenes. In general both **monocular and binocular edges seem to enhance the perception of stereoscopic slant**.

CHAPTER 4

Surface perception user study

For the experiment 19 participants where invited. Each of which had to pass an initial test for stereo vision called TNO test. Two of the participants did not have good enough stereo vision and were therefore excluded from the experiment beforehand. 7 of the remaining 17 observers (8 male and 9 female) had normal vision and 8 were visually impaired but their vision was corrected either with glasses or contact lenses. The remaining 2 participants were slightly short sighted, but could see properly far beyond the 60 cm distance to the display. The average age was 22.88 years with a range of 16 to 38 years.

The user study was conducted using the volume visualization software ”Volumeshop”, which was developed by Stefan Bruckner in 2005 [BG05]. For this, a preexisting plug-in was reconfigured to facilitate the experiments. The original gauge figure task plug-in for Volumeshop was developed by several researchers at the computer graphics institute including Matthias Bernhard and has since been used in experiments such as those by Solteszova et al [STPV12]. Although Volumeshop was designed for volumetric data models, the gauge figure experiment was conducted using geometric data models for easier calculations.

4.1 Stimuli

4.1.1 Objects

In the experiments the participants had to perform the gauge figure task on four convex objects for a discrete set of ground-truth normals, i.e. controlled variations of slant and tilt. Figure 4.1 shows the four used objects. On one hand, the experiments should use a ”best case” set of objects with a minimum of shape ambiguity. This was achieved by using primitive objects for which it can be assumed that they are well-known by all users. Furthermore, these objects were displayed with textures, specular and diffuse shading as well as motion cues, i.e. a short animation where the objects were rotated 360 degrees in 2000ms.

On the other hand, it is important to test extreme cases with increased levels of complexity and less cues. For this, a worst-case condition was included using an organic randomly deformed sphere (in the experiments simply referred to as ”Blob”) where shape information is only provided by intensity variations caused by diffuse Lambertian shading.

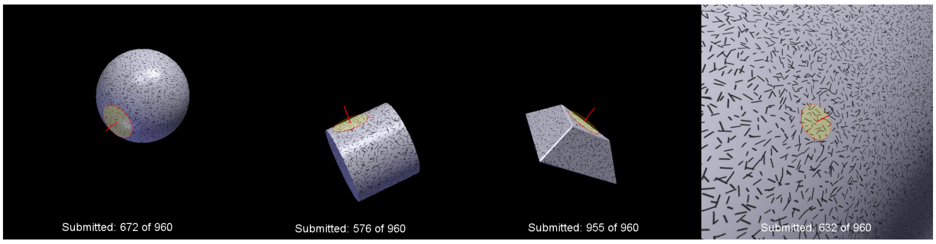

Figure 4.1: The four objects used for this user study: sphere, cylinder, simplex stump and blob. [vK04]

Each of the four objects represents a different surface category with increasing shape ambiguity.

- Sphere: represents a two-dimensionally curved surface. The tangent plane at any point touches the surface at only one point.
- Cylinder: represents a one-dimensionally curved surface. The gauge figure has to be estimated on the curved surface, so that any tangent plane intersects with the curved surface at a 3D line.
- Simplex Stump: represents a plain surface. A cube would also satisfy the requirement of a plane surface, but a simplex stumps avoids trivial cues, such as congruences between gauge figure features and shape features, which correlate with the ground truth normals. The edges were chamfered to enhance shape contours. In addition, the sample points were only estimated at the top triangle. The tangent plane of any sample point in the top triangle coincides with the triangle itself.
- Blob: represents an organic surface containing smooth random curvatures, similar to the objects used by O’Shea et al. [OBA08]. For this object the same material was used as with the other objects, but specular reflections were disabled. This was necessary because of the general lack of visual cues. Any specular highlight with a large variance would otherwise be interpreted as a bump by the participants, even though the actual geometry might be rather flat. In contrast to the other objects, the blob is never shown in full size. It is always zoomed in to create the effect of an uneven terrain with smooth bumps but a general convex appearance.

4.1.2 Lighting and Materials

The material properties for the experiments had to provide strong visual cues in order to facilitate proper normal estimates. This is mainly achieved by using Lambertian shading and specular highlights for the sphere, the cylinder and the simplex stump. As explained in previous work Lambertian shading along with highlights enhance the precision and accuracy of surface normal estimates. The blob was also shaded using Lambertian shading, but without the specular highlights, since they may produce the impression of smooth bumps in flat ambiguous surfaces.

Constant lighting conditions were used for all objects and viewing conditions. The objects were illuminated with three-point lighting, which is a standard method to light single objects both in real and virtual scenes such as in theater, film, photography and CGI. This technique uses three lights and is illustrated in Figure 4.2. The first light is the "key light", which is usually the strongest light and sets the general lighting of the scene. The second light is called "fill light" and helps to fill shadows that the main light might cast. The last light is called "backlight" or "rimlight" and is used to create a contour and separation from the background.

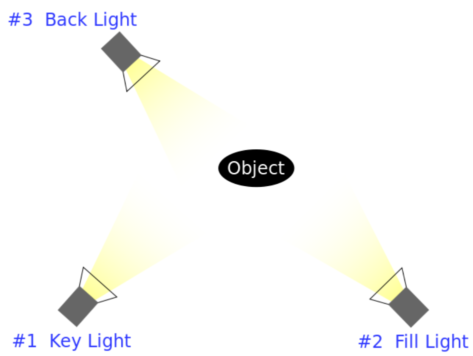

Figure 4.2: Standard three-point-light setup to illuminate an object. (Image courtesy Wikimedia Commons)

In contrast to the marble 3D texture used by Solteszova et al. [STPV12], the four objects were textured using Poisson distributed random line segments on white background. Marble textures generally have texture gradients, which might influence the samples in a negative way, similar to specular highlights looking like smooth bumps. This "line texture" also allows for a better fusion of the two slightly different images in the stereoscopic viewing conditions, since it is easier for the brain to fuse discrete objects with high contrast to the background.

To generate such a tileable texture a texture generator was implemented in Matlab. The generator can create random Poisson distributed or uniformly distributed line textures, Figure 4.3. The Poisson distribution of lines was necessary to guarantee an even yet random distribution of lines across the entire object and was implemented using the blue noise sampling algorithm designed by Robert Bridson [Bri07]. If the lines are distributed uniformly there is a tendency to create clusters and void areas in the texture, Figure 4.4. The lines are created by using a modified version of the Bresenham's line algorithm [Bre65]. Several parameters can be adjusted to generate the most suitable texture for a given task, e. g. image size, line density, line width, minimum and maximum line length as well as background and line color. The generated images can be filtered by a Gaussian blur with adjustable sigma and kernel size. This reduces aliasing artifacts in the final texture.

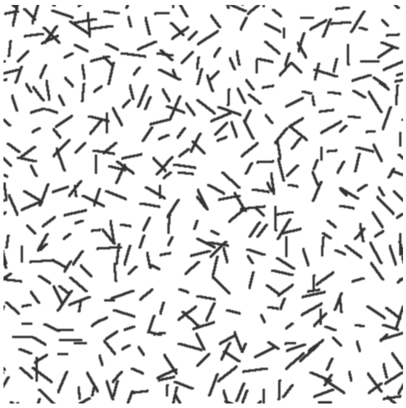

Figure 4.3: Example of a Poisson distributed, random, tileable, line texture.

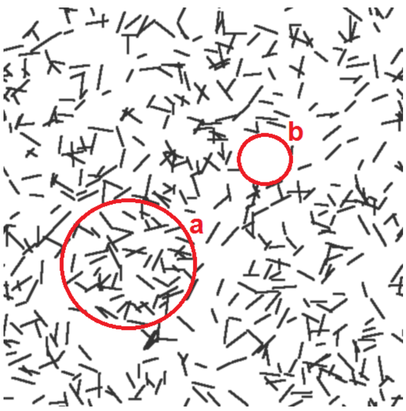

Figure 4.4: Example of a uniformly distributed, random, tileable, line texture. Notice that region ”a” is a cluster of lines and in region ”b” are no lines at all. This can be avoided by using Poisson disc sampling.

Furthermore, the generator is able to not only generate tileable square textures but also spherically mapped images as shown in Figure 4.5. This was necessary in the beginning of the user study design to properly map the textures onto the sphere and the blob. Both of those objects are not easily mapped because the texture is always distorted. When mapping 2D textures onto a sphere-like object there are several desirable but contradicting aspects.

- Preserving angles
- Preserving shape locally

- Preserving area
- Preserving distance

In fact, it is impossible to design a mapping that is both equal in area and preserves the local shape. Furthermore, preserving the distances between all possible mapping points is not possible either.

The initial idea was to counteract and limit these mapping-distortions by distorting the texture beforehand. This approach worked, but in the end I found that Autodesk Maya is able to map UV coordinates spherically, which renders the initial distortion approach somewhat useless. By mapping all objects, including the sphere and blob, with UV coordinates in Autodesk Maya, it was possible to use the same generated texture for all objects, which accelerated the loading time for the objects in the experiments.

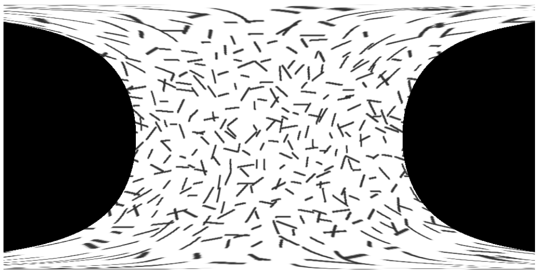

Figure 4.5: Example of a spherically mapped line texture.

4.2 Experiment Setup and Viewing Conditions

The experiment setup was rather constrained to avoid influencing factors as much as possible and to configure the stereoscopic 3D system specifically for each user. For the experiment a monitor capable of displaying at the speed required for shutter glasses (120 Hertz) was placed on a table. The participants were placed in front of it with a distance of 60 cm between their eyes and the screen. The users rested their heads on a chin rest, which was adjusted for each participant so that their head was at a height where their eyes were centered perfectly in regards to the screen. Figure 4.6 shows an illustration of this setup.

Five viewing conditions were investigated in this user study:

1. V1: monocular viewing using only the dominant eye while the other eye was covered by an eye patch
2. V2G1S1: binocular viewing with both monoscopic gauge figure and stimuli
3. V2G2S1: binocular viewing with stereoscopic gauge figure and monoscopic stimuli
4. V2G1S2: binocular viewing with monoscopic gauge figure and stereoscopic stimuli

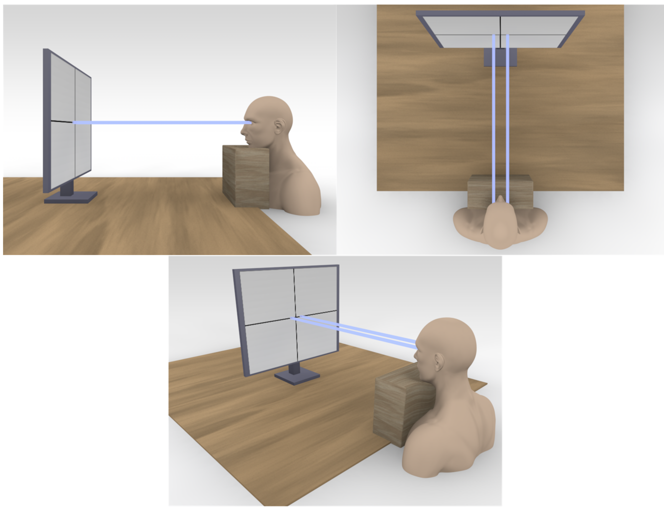

Figure 4.6: Basic experiment setup.

5. V2G2S2: binocular viewing with both stereoscopic gauge figure and stimuli

In condition V1 only the dominant eye was used rather than the non-dominant, because of an interesting aspect of human vision called ocular dominance or ”eye dominance”. The term describes the tendency to prefer one eye over the other. Studies show that the majority of people (65.81%) are right-eyed and the eyedness is not influenced by sex or handedness [DL76], but the dominant eye is the one that is generally used for precise positional estimation [EASGJ05]. It was important to take this into account in the experiment design to avoid possible errors introduced by inaccuracies when using the non-dominant eye.

Conditions V1 (monocular) and V2G1S1 were included in order to test the hypothesis that monoscopic viewing improves surface perception. There is no depth-cue conflict between the shown 3D cues and the flat screen observed by participants in the monoscopic viewing condition since there is only one eye seeing one image rather than two eyes seeing slightly different images. The condition V2G1S1 serves two main reasons. First, it is the control condition for possible improvements with the pure monoscopic condition V1. Second, it is used to reconfirm previous findings of underestimation of surface normals.

Condition V2G1S2 and V2G2S1 are used to confirm or falsify the hypotheses that the gauge figure task itself is the source for the systematic underestimation of surface slants. Unfortunately, while researching the previous work I found that the experiment setups are often not explained properly. Therefore, it is often unclear whether a stereoscopic gauge figure or a simple 2D one with stereoscopic stimuli was used.

Lastly, condition V2G2S2 with full stereoscopic vision is used to eradicate influences of depth-cue conflicts as much as possible. The ideal gauge figure task would work on real objects, but since evaluating such estimates are prone to error the virtual alternative as used in this user study is the only practical way of conducting such experiments in a reliable way. The full stereoscopic condition simulates the estimation of real world objects by using exactly calibrated stereoscopic viewing parameters for every participant, so that the two virtual cameras are located at the same virtual position as the real position of the participant's eyes in front of the monitor. This was mainly achieved by measuring the participant's interpupillary distance and by fixating the participant's head with the chin rest. Furthermore, the four stimuli objects were placed within the plane of zero parallax. To be more precise, the point of intersection between an object's surface and the stick of the gauge figure was placed in the plane of zero parallax in the middle of the screen.

4.3 Sampling points and participants

In this user study the ground truth normals were compared with the user estimates, which had not been possible with the original gauge figure task developed by Koenderink et al. [KvDK94]. Since it is not feasible to sample every possible sample point on a given surface it was necessary to identify a manageable set of sample points that cover the most important areas on the half sphere created by all possible slants and tilts.

Therefore, the sampling half sphere was divided into discrete areas by using specific slants and tilts. 6 different slants with 15 degrees step size were used: 5, 20, 35, 50, 65 and 80 degrees. Slants of 0 and 90 degrees are special cases and were intentionally excluded due to their possible negative effects in the analysis phase.

With a slant of 0 degrees the stick of the gauge figure points directly at the viewer, looking like a point surrounded by a perfect circle, i. e. the undistorted gauge figure disc. Since the slant is zero, the tilt does not exist, because it is a pole in the mathematical sense of the word. Any deviation in any direction would cause problems with evaluation. With slants (even slightly) bigger than 0 there is always an associated tilt. And because these tiny deviations are rather random, the sampling results would be altered in a negative way. In other words, a slant of 0 degrees was excluded to counteract the sampling of tilts that are not defined.

A slant of 90 degrees was also excluded because the gauge figure's movement is constrained by the sampling half sphere. Sampling a point with a slant of 90 degrees would simply mean to rotate the gauge figure to the side until it does not move any more. At the border of the sampling half sphere, i. e. at 90 degrees slant, the gauge figure disc also looks like a stick rather than an ellipse, which would make estimating such a slant even easier.

The possible tilts of 360 degrees were divided into 8 disjoint, non-empty bins, each with a range of 45 degrees. Within each bin a random tilt was chosen in order to provide a sample point for each slant-tilt combination.

Since there are 4 objects to be estimated under 5 different viewing conditions (V1, V2G1S1, V2G2S1, V2G1S2, V2G2S2) and for each combination 6 slants and 8 tilts are to be estimated, each participant had to estimate a total of 960 points.

In the experiments one location on the objects surface was randomly selected and then the object was rotated until it aligned with one of the predefined sample points. Moreover, the object was translated in such a way that the gauge figure's position in screen space was in the center of the screen, where maximum accuracy and lowest amount of distortions caused by perspective rendering and stereoscopic viewing were assumed. Depending on the object certain areas are not applicable as surface sampling point, e. g. the sides of the simplex stump or the top and bottom circle of the cylinder. These areas, as well as points that cause the object when aligned to be cut off, were not used.

During the user study I collected the estimates from 17 paid participants. Since there were 17 participants and each of them estimated 960 points, 16320 surface estimates were collected in total and used for the analysis.

4.4 User study procedure

The 960 sample points per participants were separated into two blocks by conditions, one mono block where the stimuli are in mono with 576 samples (V1, V2G1S1, V2G2S1) and one stereo block where the stimuli are in stereo with 384 samples (V2G1S2, V2G2S2). The participants were invited for one session for each block with an average duration of 1 to 2 hours per session. Between the two blocks there had to be a break of at least one day in order to reduce the learning effect of the objects and tiredness of the users. The entire experiment can be divided into 11 separate steps, which are explained in detail below.

1. **Declaration of consent**: First, the participants were asked to sign a declaration of consent and had to read a short introduction about the general procedure of the experiment and the gauge figure task. The participants were told that they could end their participation at any time, especially if they felt uncomfortable due to the active shutter 3D system. Any details about the scientific goals of the experiments were kept secret until the end of the second block in order to prevent biases.

2. **Practicing the gauge figure task**: Second, sample interactions with the interface of Volumeshop were shown to the participants. The users were asked to practice estimating surface normals of a random object on a by-standing laptop with the same user interface as in the actual experiments until they felt comfortable enough with the gauge figure task. In addition, the participant were handed three real world models of the objects they later had to estimate (sphere, simplex stump and cylinder). The objects had the same size as the displayed objects in the actual gauge figure task. The idea behind this was to reduce top-down ambiguity in the three simple objects and to make the objects well-known by the participants.

3. **Stereo vision test**: After practicing the interaction, the observers had to undergo a stereo vision test called TNO-Test. This test uses the anaglyph 3D stereo technique. The participants were shown 5 images. Each of which contained several objects that are only visible when one is able to see in 3D, i. e. the test person is able use stereoscopic fusion. With each image the difficulty of seeing the hidden objects increased. If a participant was not able to identify the hidden images in at least 3 out of the 5 test images, the person was excluded from the experiments since their estimates for the conditions requiring stereo vision would be insufficient.

4. **Questionnaire**: In this step the participants had to answer some questions regarding things like their eye-sight, previous experience with computers and 3D systems as well as demographic data such as age and gender. The results of these questionnaires are discussed in Chapter 5.2. In addition, this was the stage were it was decided whether the participant would start with the mono or with the stereo block. The goal was to achieve an equal amount of users starting with the mono and with the stereo block. In a pilot run of the experiment it was found that the stereo block might cause faster object learning than the mono block, i. e. users performed better if they started with the stereo block. This procedure of alternating between mono and stereo block at the beginning might help to even out the influence of fast object learning in the stereo block versus slow object learning in the mono block.

5. **Measurements and experiment setup**: The next step was measuring the interpupillary distance of the participant to properly configure the stereo system. Furthermore, the chin rest was adjusted in order to center the user's eyes in regards to the monitor and to adjust the distance to the screen to 60 cm. Lastly, the position and orientation of the monitor was checked and corrected if necessary.

6. **3D shutter system setup**: At this point in the user study procedure it was necessary to run the NVIDIA stereo setup wizard in order to make sure that it worked properly in the experiment.

7. **Volumeshop check**: Before the experiment could start, a last check of settings in Volumeshop was performed. This included checking the interpupillary distance in the stereo setup, field of view, object sizes, parallax and sample repetitions.

8. **Aperture**: At this stage of the experiment all configurations were completed. A physical aperture was positioned in front of the screen and the lights in the room were turned off to achieve almost total darkness. The aperture was necessary to prevent visual conflicts at the border of the monitor when the blob was viewed stereoscopically. I used a physical aperture rather than a virtual one to guarantee that the participants would not mistake the edges of the aperture for the silhouette of the blob.

9. **Gauge figure task**: The participants estimated the surface normals. In the stereo block they had to complete 2 and in the mono block they had to complete 3 viewing conditions. During the task the participants were kept in a conversation and between conditions a short break was held to reduce tiredness.

10. **Qualitative assessment**: After each viewing condition the participants were asked a few qualitative questions, including whether their focus was mainly on the gauge figure disc or stick, whether they looked locally or globally at the object and what object they thought to be the hardest and the easiest to estimate. The results of these questions are discussed in Chapter 5.2.

11. **Financial compensation**: If the participant successfully estimated an entire block, they received a financial compensation of 15,00 Euro per block, i. e. 30,00 Euros for both blocks.

Steps 1 to 4 were performed only once at the beginning of the first block. The following steps 5 to 11 were carried out in the second block as well (except measuring the interpupillary distance).

CHAPTER 5

Results and Discussion

5.1 Results

In Figure 5.1 the mean slant error for each estimated tilt-slant combination is visualized. As expected, the results confirm that a clear slant underestimation (blue) occurs in the mono conditions $V1$ and $V2G1S1$ for the blob, whereas simple and known objects reduce slant underestimation across all conditions. The minimal error occurred in the full stereo condition $V2G2S2$ (almost white in the visualization). In the statistical analysis the slant errors across all tilts within each hemispherical sector were aggregated using the median, which is more robust than the mean. For the statistical analysis an ANOVA was used, Chapter 7 explains how the results of such an analysis may be interpreted.

The ANOVA investigating the effects of the three well-known objects compared to the blob in modality-consistent viewing conditions confirms that well-known objects lead to a significantly lower systematic errors in the mono-viewing conditions ($V1, V2G1S1$) compared to the blob with a medium to large effect-size ($F(1,732)$=102, p<0.001, $\hat{\eta}_p^2$=0.12). However, the difference between the three well-known objects and the blob fully disappears when the stimulus is presented in stereo ($V2G2S2$, $F(1,395)$=102, p=0.8, $\hat{\omega}_p^2$=0.00).

Comparing the all-mono and all-stereo conditions revealed that there is a strong effect for the blob ($F(1,256)$=86, p<0.001, $\hat{\eta}_p^2$=0.25). As expected, this effect is quite reduced for well-known objects ($F(1,871)$=18, p<0.001, $\hat{\eta}_p^2$=0.02). However, a deeper analysis showed no effect of stereo-viewing a sphere ($F(1,280)$=0.4, p=0.50, $\hat{\eta}_p^2$=0.00), whereas a small-to-medium effect for the other well-defined objects ($F(1,579)$=27, p<0.001, $\hat{\eta}_p^2$=0.04) is present. This is a clear indication that the sphere provides the most accurate and reliable reference object to evaluate probing accuracy, whereas the other well-known objects introduce a limited yet significant amount of perceptual ambiguity. In addition to the sphere being the most accurately estimated object it also shows a unique symmetrical wave-shaped pattern of under- and overestimation, which is not present for the other objects.

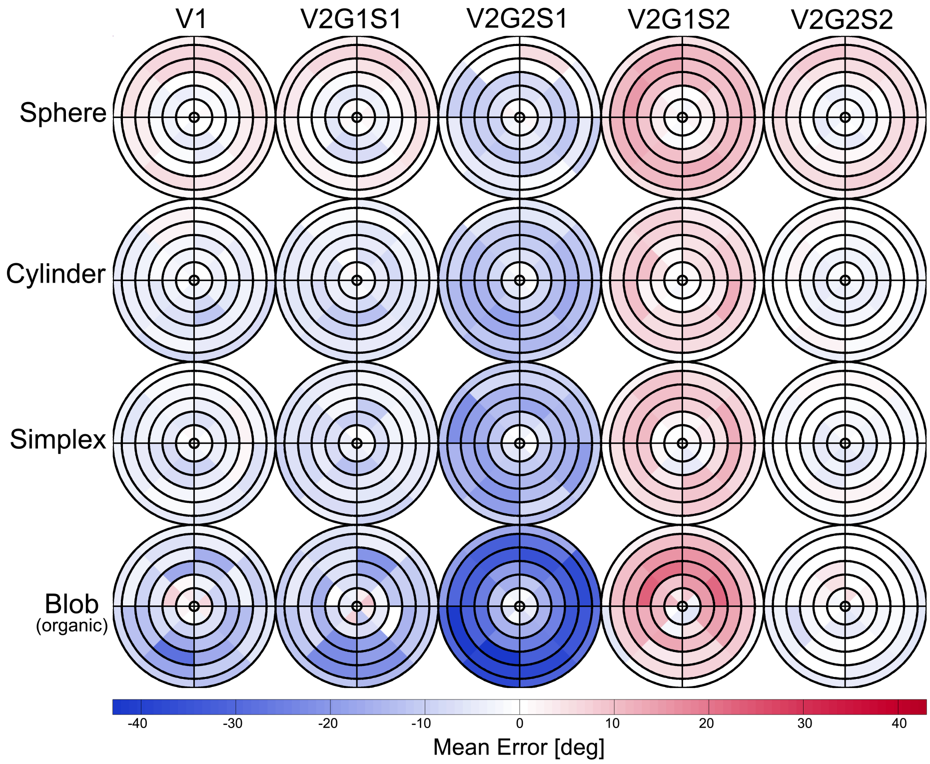

Figure 5.1: Mean slant error across the sampled hemisphere for all participants.

In addition, there is a small effect of using one eye instead of two eyes for the gauge figure task on a mono display ($V1$ vs. $V2G1S1$). While there was a significant effect for the blob ($F(1,155)$=17.0, p<0.0001, $\hat{\eta}_p^2$=0.10), the effect of mono-eye viewing is rather negligible with the three well-known objects as stimulus ($F(1,565)$=3.9, p=0.048, $\hat{\eta}_p^2$=0.01). Therefore blindfolding one eye is a reasonable strategy to improve accuracy, though the differences are very subtle on an absolute scale, especially for simple and known objects. This is also reflected in the slant accuracy curves in Figure 5.2. The curves for the monocular case ($V1$) and the binocular case without stereoscopic vision ($V2G1S1$) look very similar, but there is marginally less deviation from the ground truth slants in the monocular case.

Furthermore, the analysis also revealed an enormous stimulus slant underestimation for mono stimuli if the gauge figure was shown in stereo ($V2G2S1$, indicated by blue color in Figure 5.1 and also clearly visible in 5.2). When comparing slant errors from $V2G2S1$ to those collected in the modality-consistent condition $V2G1S1$ this effect is rather extreme for the blob ($F(1,240)$=362.7, p<0.001, $\hat{\eta}_p^2$=0.60) and also surprisingly large for the three well-known objects ($F(1,855)$=362.7, p<0.001, $\hat{\eta}_p^2$=0.19). Likewise, a pronounced slant overestimation (indicated by red color in Figure 5.1 and also visualized in 5.2) occurs when the gauge figure is in mono but the stimulus in stereo ($V2G1S2$),

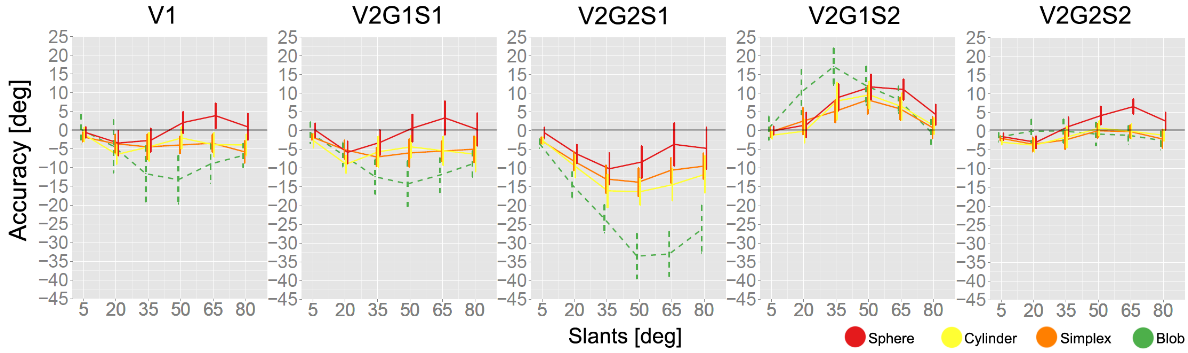

Figure 5.2: Slant accuracy (mean error) for all objects and all viewing conditions.

which has a large effect, though lower than in the inverse case, when the object is ambiguous ($F(1,179)$=131.3, p<0.001, $\hat{\eta}_p^2$=0.42) and notable for the well-known objects ($F(1,593)$=190.6, p<0.001, $\hat{\eta}_p^2$=0.25).

These results may have two explanations: First, the gauge figure task's accuracy is affected by an inconsistency in the viewing modality, where probing is biased by an interaction between stereo and mono. Second, the slant of the gauge figure's disc seen in mono is systematically underestimated and consequently the measurement overestimates the perceived slant of the stimulus. To shed more light on these issues, three statistical models for slant accuracy were developed by Matthias Bernhard. The three explanatory models are:

1. Model 0: The null model assumes that all errors are the result of a skewed perception of scene depth.

2. Model 1: This model speculates that in the worst-case the gauge figure task introduces a wave-like pattern, but apart from that all effects are perceptual.

3. Model 2: This model assumes that the slant of the mono gauge figure is systematically underestimated and yields an overestimation of the perceived slants in the probes.

The models are explained in detail in the coauthored technical report [BWP+15]. For an assessment and a more in depth discussion of the results refer to Chapter 5.3.

5.2 Qualitative Assessment

All 17 participants had previous experience with 3D stereoscopic vision in either cinematic movies or at home, e.g. on their TV or PC. 13 said that they experienced the 3D effect as "strong" and the other 4 as "moderately strong".

There were three main groups of participants: students (8), pupils (5) and researchers (4). Two participants from the pupil group with very little day to day practice on PCs initially had problems with the gauge figure task. Both started with the monoscopic block and said that they "could not imagine the displayed objects as 3D". Only after about 50 to 70 samples they mentally "understood" the objects and performed properly. Another special case were two of the four participants from the researchers-group. Their results were substantially different from the others, probably due to their previous work with the gauge figure task. The errors for the two of them in the monocular configuration were contradictory to each other, one underestimated the object's surface normals strongly and the other one overestimated them. For one of them, the error of slant estimates in the other four configurations were significantly smaller than the average of all participants, possibly from the many previously performed gauge figure tasks in other experiments.

Furthermore, the results improved if the participants performed the two configurations with stereoscopic stimuli before the three monoscopic cases. These special cases and the improved results with previous exposure the stimuli in stereoscopic vision indicate that there might not necessarily be a learning effect for the gauge figure task itself, but there seems to be a strong learning effect for the displayed objects.

As expected, in general there was a strong consensus among all observers that the stimuli under stereoscopic vision were much easier to estimate. For most observers (11) the easiest object to estimate was the sphere, the hardest the organic shape (15 participants). The organic shape was harder to estimate for smaller slants and easier for higher ones. This makes sense, since the organic shape for small slants looks like a flat surface without much information about the object's surface orientation, because there is zero to a very small texture gradient. Participants described that they felt "lost" in the flat surface with no apparent surface direction. For higher slants the texture gradient is stronger and the participants had stronger visual cues for estimating the object's surface.

These statements are reflected in the curves for the slant precision (standard deviation) of the organic shape in three cases: $V1$, the mono-consistent case $V2G1G1$ and the stereo-stimuli-case with the gauge figure in mono $V2G1S2$, Figure 5.3. For higher slants the participants felt more confident with lower deviations in their actual results. In the two cases with stereoscopic gauge figures it seems as if the participants used other cues such as disparity to estimate the organic shape, hence the curves of slant deviations in these two cases do not show this negative correlation as much. This suggests that for unknown geometry a stereoscopic gauge figure, rather than a monoscopic one, seems to improve precision of estimates.

8 participants reported that they perceived more depth in the monocular case in comparison to the binocular case without stereoscopic vision. 6 observers said that they experienced no change in depth. This confirms the findings of Koenderink et al. [KvDK94] and is slightly reflected in our results as mentioned in the previous chapter.

In the binocular configuration where the stimuli were presented in mono and the gauge figure in stereo, 5 participants reported that the gauge figure seemed to hover or float above the object. This uncoupling of gauge figure and stimuli might be one possible reason for the strong underestimation in the results, since the uncoupling introduces a perceptual conflict.

Most participants (11) used both the disc and the needle of the gauge figure to

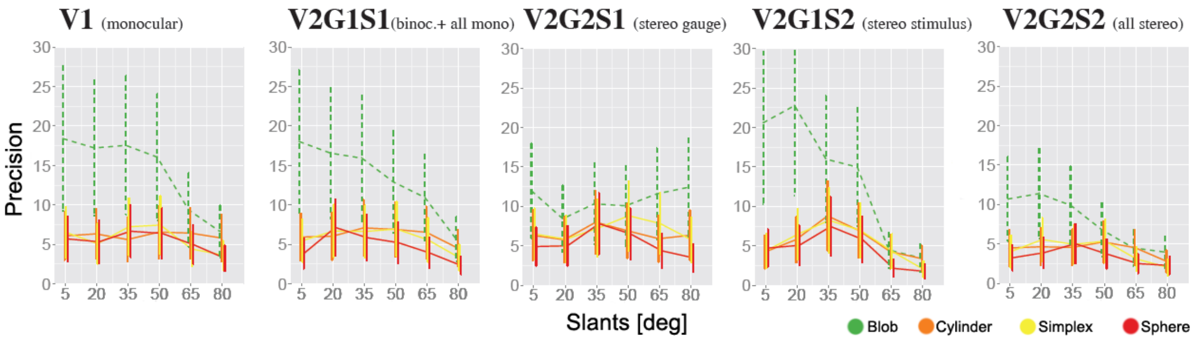

Figure 5.3: Slant precision (standard deviation) for all objects and all viewing conditions. In general, the higher the slant for the unknown object (blob) the more precise the estimates.

estimate the surface slant. Usually the observers used one of the two to set the general direction and afterwards adjusted or confirmed their initial gauge figure setting using the other one. Participants that mostly used only the disc (2) or the needle (4) for their estimates preferred the needle. The concept of estimating a "surface normal" rather than a "tangent plane" seems to be easier for estimating the surface orientation even though both concepts are mathematically equivalent.

For higher slants for the sphere some participants (13) used the outer rim of the sphere to adjust the gauge figure for the first few estimates. They tended to align the rim of the gauge figure disc with the rim of the sphere in 2D as shown in Figure 5.4. As the experiment continued they all realized that this technique is "wrong" and that they should use a higher slant. This systematic and somewhat unintuitive behavior in the beginning of some experiments might influence our results. Therefore, the samples for higher slants for the sphere object might not be as reliable as the others.

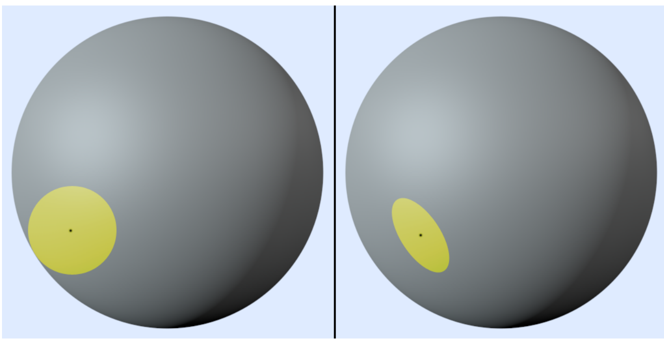

Figure 5.4: The left sketch shows how some participants aligned the rims of the sphere and gauge figure discs in 2D rather than estimating the 3D surface. The right sketch shows a better estimate of the sample point.

The participants were also asked whether they generally looked at the area around the gauge figure or at the presented object globally to estimate the surface orientation. Most reported that they looked both ways (8 participants), usually using one to set the general direction and the other one for final adjustments. Participants that used mainly one way to look where almost equally divided into ”local” (5) and ”global” (4) observers.

During the stereo part of experiment the participants were asked whether the sphere looked normal, elongated in view-direction causing an egg-shape, or whether the sphere was squished. All observers answered that they saw a perfect sphere, indicating that the stereo setup was accurate enough for our measurements. This contradicts one possible and negative interpretation of the results that an incorrect stereoscopic setup might be the cause of the under- and overestimation in slant estimates.

Lastly, almost all participants reported that the task itself is rather tiring. To counteract this expected effect when probing a lot of sample points short brakes were held between each configuration and the participants were kept active by talking to them. Nevertheless, this tiring might still have had a negative effect on some the gathered sample points increasing the deviations slightly.

5.3 Discussion

As expected, there was only a very small difference between the all-mono and the all-stereo conditions for all well-known objects, which was practically zero in effect-size for the sphere. In contrast to this, the ambiguous blob showed a strong slant underestimation for the all-mono condition, but the underestimation was eliminated in the all-stereo condition. This could indicate that the perception of the three well-known objects was already optimal in the all-mono condition and could not be improved any further - under the assumption of fully precise gauge figure measurements.

However, the strong slant overestimation in the $V2G1S2$-condition, i.e. stereo stimulus with a mono gauge figure, compared to the all-mono condition for all objects shows that surface perception is sensitive to changes in the viewing condition - even for well-known objects. Therefore, one cannot expect that all perceptual factors are eliminated by cognition, even for well-defined objects with strong monoscopic depth cues.

The largest error between ground truth slant and probed slant was observed for the $V2G2S1$ condition, where the gauge figure was rendered in stereo and the stimulus in mono. The second largest error was measured in the $V2G1S2$ condition with a mono gauge figure and a stereo stimulus. This shows that the perceptual understanding of the stereo presentation does not align with the mono presentation, at least when they are presented simultaneously. One reason could be the disparity contrast between the gauge figure and the stimulus. Another explanation could be that participants fail to accurately ”communicate” a slant perceived with a gauge figure presented in mono, and vice versa.

Given the fact that well-known objects like the sphere can be estimated with very high accuracy in stimulus-consistent conditions, an interpretation using the following factors might be the most reasonable explanation. The observations may be best explained by

1. disparity contrast

2. unknown effects occurring when a participant performs an intra-modal mapping between mono and stereo

3. a perceptual mismatch between mono and stereo conditions, which might even be partially a result of inaccuracies in the experimental setup

4. the sum of these three factors

Overall, even if the slant of surfaces on well-known objects and the gauge figure's disc are both underestimated in mono, at least an upper bound for the gauge figure task probing error can be defined.

CHAPTER 6

Conclusion and Future Work

This bachelor thesis was the first attempt to investigate the error introduced by the gauge figure task to probe human surface perception. The result was that slants estimated with the gauge figure task are accurate within the modality they are taken, i. e. when a mono stimulus is probed with the gauge figure shown in mono as well, or vice versa.

When the gauge figure and stimulus are presented in stereo ($V2G2S2$) the observed systematic deviation of estimated slants from the ground truth slants, which is the sum of all perceptual and probing errors, is almost optimal within $\pm 3°$. This still holds true for ambiguous stimuli like the blob. In the mono-consistent viewing conditions ($V1$, $V2G1S1$) the accuracy is low for ambiguous surfaces like the blob, but good for well-known stimuli. The minimum error was observed for the sphere and had a symmetrical wave-shaped pattern, which was bounded within a range of $\pm 5°$.

In regards to slant precision, i.e. slant standard deviation, the results suggest that at least for unknown geometry a stereoscopic gauge figure, rather than a monoscopic one, seems to improve precision of estimates. This, along with the higher accuracy errors in modality inconsistent conditions, justifies our study approach to investigate the effects of stereoscopy for all four possible stimuli and gauge figure combinations.

There is still an open question whether there is an accurate correspondence mapping between estimates taken in mono to ones taken in stereo. In the experiment mixed viewing conditions were included where the stimuli are seen in stereo and the gauge figure was displayed in mono, or vice versa. These two special conditions require more than just specifying a gauge figure orientation, which perceptually aligns with the orientation of the probed surface. Rather, the participants need to perceive the surface orientation in one modality and communicate their understanding of this orientation in another. The experiment showed a clear deviation between the gauge figure probes taken in stereo and the stimulus presented in mono, which yields slant underestimation ($> -10°$), and vice versa, which yields slant overestimation ($< 12°$).

This deviation may be a result of a mismatch of the perceptual depth-scale between both modalities, or it is the result of an interaction of the simultaneous presentation of both modalities which amplifies or evokes a contrast (e.g., disparity contrast) in depth perception between gauge figure and stimuli. In any case, the difference is that monoscopic perception is underestimated in slant relative to stereoscopic perception. This means that the results give an upper bound for the accuracy of the gauge figure task, i. e. the worst-case probing in mono overestimates the perceived slant with a maximum error of $-10°$ (for the sphere). This result probably requires some reproductions in order to provide sufficient evidence of this result. In order to get results, which may allow to derive a more solid explanation of this effect, it would be useful to systematically vary other independent variables such as viewpoint, probing location and stereo parameters.

CHAPTER 7

Appendix A: ANOVA

For the analysis an ANOVA was used. The ANOVA provides several variables for e.g. indicating the significance. The F-value or F-ratio (= between-group variability / within-group variability) is one of these variables and is about 1 when the average difference between groups is similar to that within groups. As the average difference between groups becomes greater than that within groups, the F-ratio becomes larger than 1. x and y in $F(x, y)$ indicate the degrees of freedom for variance between groups ($x =$ number of groups $- 1$) and within groups ($y =$ total number of observations $-$ number of groups).

Another variable is the p-value. To obtain a p-value, the F-ratio can be tested against the F-distribution of a random variable with the degrees of freedom associated with the numerator and denominator of the F-ratio. The p-value is the probability of getting that F-ratio or a greater one. Larger F-ratios give smaller p-values. A small p-value (typically < 0.05) indicates strong evidence against the null hypothesis, so the null hypothesis (= assuming there is no significant effect) is rejected. Therefore rejecting the null hypothesis means there is a significant effect.

The variable $\hat{\eta}_p^2$ means partial Eta-squared and indicates the effect size. It is a quantitative measure of the strength of a phenomenon and allows comparisons of an effect between different studies or experiments. In general $\hat{\eta}_p^2 = 0.01$ indicates a small, $\hat{\eta}_p^2 = 0.06$ a medium and $\hat{\eta}_p^2 = 0.14$ a strong effect.

The variable $\hat{\omega}_p^2$ is similar to $\hat{\eta}_p^2$ and is also a measurement of effect size. It introduces less distortions than partial Eta-squared since it includes the number of groups in the calculation. The difference between $\hat{\omega}_p^2$ and $\hat{\eta}_p^2$ is usually small and is even more reduced with larger sample size [SR13].

Bibliography

[BG05] Stefan Bruckner and Meister Eduard Gröller. Volumeshop: An interactive system for direct volume illustration. In H. Rushmeier C. T. Silva, E. Gröller, editor, *Proceedings of IEEE Visualization 2005*, pages 671–678, October 2005.

[BKY97] P.N. Belhumeur, D. Kriegman, and A.L. Yuille. The bas-relief ambiguity. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 1060–1066, Jun 1997.

[BM88] Heinrich H. Buelthoff and Hanspeter A. Mallot. Integration of depth modules: stereo and shading. *J. Opt. Soc. Am. A*, 5(10):1749–1758, Oct 1988.

[Bre65] J.E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.

[Bri07] Robert Bridson. Fast poisson disk sampling in arbitrary dimensions. In *ACM SIGGRAPH 2007 Sketches*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.

[BWP+15] Matthias Bernhard, Manuela Waldner, Pascal Plank, Veronika Solteszova, and Ivan Viola. The accuracy of gauge-figure tasks in mono and stereo displays. Technical report, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11, A-1040 Vienna, Austria, September 2015.

[Cra] Reallusion Crazytalk. The concepts of 3d stereo vision. `https://www.reallusion.com/de/iclone/help/iclone5/PRO/09_3D_Vision/The_Concepts_of_Stereo_Vision.htm`. Accessed: 18 Junel 2015.

[CSD+09] Forrester Cole, Kevin Sanik, Doug DeCarlo, Adam Finkelstein, Thomas Funkhouser, Szymon Rusinkiewicz, and Manish Singh. How well do line drawings depict shape? *ACM Trans. Graph.*, 28(3):28:1–28:9, July 2009.

[CSR84] B. Cassin, S. Solomon, and M.L. Rubin. *Dictionary of eye terminology.* Triad Pub. Co., 1984.

[DHEN95] Erik De Haan, Roderik G.F. Erens, and Andre J. Noest. Shape from shaded random surfaces. *Vision Research*, 35(21):2985 – 3001, 1995.

[DL76] Chaurasia B. D. and Mathur B. B. L. Eyedness. *Cells Tissues Organs*, 96(2):301 – 305, 1976.

[EASGJ05] Walter H. Ehrenstein, Birgit E. Arnold-Schulz-Gahmen, and Wolfgang Jaschinski. Eye preference within the context of binocular functions. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 243(9):926–932, 2005.

[FL13] Arthur Faisman and Michael S. Langer. How does lighting direction affect shape perception of glossy and matte surfaces? In *Proceedings of the ACM Symposium on Applied Perception*, SAP '13, pages 9–14, New York, NY, USA, 2013. ACM.

[GBC+88] C. C. Gordon, B. Bradtmiller, C. E. Clauser, T. Churchill, J. T. McConville, I. Tebbetts, and R. A. Walker. Anthropometric survey of u.s. army personnel: Methods and summary statistics 1988, 1988.

[HHR95] Ian P. Howard, Aan P. Howard, and Brian J. Rogers. *Binocular Vision and Stereopsis (Oxford Psychology Series).* Oxford University Press Inc, 1995.

[JP94] Alan Johnston and Peter J. Passmore. Independent encoding of surface orientation and surface curvature. *Perception*, 23(2):169–189, 1994.

[KMS+06] Kristin Koch, Judith McLean, Ronen Segev, Michael A. Freed, Michael J. II Berry, Vijay Balasubramanian, and Peter Sterling. How much the eye tells the brain. *Current Biology*, 16:1428 – 1434, 2006.

[Koe29] Wolfgang Koehler. *Gestalt Psychology.* New York: Liveright, 1929.

[KvDC96] Jan J. Koenderink, Andrea J. van Doorn, and C. Christou. Perturbation study of shading in pictures. *Perception*, 25(9):1009–1026, 1996.

[KvDK92] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Surface perception in pictures. *Perception and Psychophysics*, 5(52):487–496, 1992.

[KvDK94] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. On so-called paradoxical monocular stereoscopy. *Perception*, 23(5):583–594, Jan 1994.

[KvDKT01] Jan J. Koenderink, Andrea J. van Doorn, Astrid M. L. Kappers, and James T. Todd. Ambiguity and the 'mental eye' in pictorial relief. *Perception.*, 30(4):431–448, Jun 2001.

[LT04] Baoxia Liu and James T Todd. Perceptual biases in the interpretation of 3d shape from shading. *Vision Research*, 44(18):2135 – 2145, 2004.

[MK96] Pascal Mamassian and Daniel Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351 – 2367, 1996.

[MT86] E Mingolla and J T Todd. Perception of solid shape from shading. *Biol. Cybern.*, 53(3):137–152, January 1986.

[Nik09] Danko Nikolic. Is synaesthesia actually ideaesthesia? an inquiry into the nature of the phenomenon. In *Proceedings of the Third International Congress on Synaesthesia*. Science and Art, Granada, Spain, 2009.

[NW07] J.Farley Norman and ElizabethY. Wiesemann. Aging and the perception of local surface orientation from optical patterns of shading and specular highlights. *Perception and Psychophysics*, 69(1):23–31, 2007.

[OBA08] James P. O'Shea, Martin S. Banks, and Maneesh Agrawala. The assumed light direction for perceiving shape from shading. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08, pages 135–142, New York, NY, USA, 2008. ACM.

[PW12] 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.

[Sch13] U. Schmidt. *Professionelle Videotechnik: Grundlagen, Filmtechnik, Fernsehtechnik, Geräte- und Studiotechnik in SD, HD, DI, 3D*. SpringerLink : Bücher. Springer Berlin Heidelberg, 2013.

[SP96] Jennifer Sun and Pietro Perona. Where is the sun? *Nature Neuroscience*, 1(3):183–184, Jul 1996.

[SPV11] Veronika Solteszova, Daniel Patel, and Ivan Viola. Chromatic shadows for improved perception. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Non-Photorealistic Animation and Rendering*, NPAR '11, pages 105–116, New York, NY, USA, 2011. ACM.

[SR13] P. Sedlmeier and F. Renkewitz. *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler*. Always learning. Pearson Deutschland, 2013.

[Ste83a] Kent A. Stevens. Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, 46(3):183–195, 1983.

[Ste83b] Kent A. Stevens. Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception and Psychophysikcs*, 33(3):241 – 250, 1983.

[STPV12] V. Solteszova, C. Turkay, M.C. Price, and I. Viola. A perceptual-statistics shading model. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2265–2274, Dec 2012.

[TI12] Nan-Ching Tai and Mehlika Inanici. Luminance contrast as depth cue: Investigation and design applications. *Computer-Aided Design and Applications*, 9(5):691–705, 2012.

[TM83] James T. Todd and Ennio Mingolla. Perception of surface curvature and direction of illumination from patterns of shading. *Journal of experimental psychology. Human perception and performance.*, 9(4):583–595, Aug 1983.

[Tod04] James T. Todd. The visual perception of 3d shape. *Trends in Cognitive Sciences*, 8(3):115–121, Mar 2004.

[VH13] Dhanraj Vishwanath and Paul B. Hibbard. Seeing in 3-d with just one eye: stereopsis without binocular vision. *Psychological science*, 24(9):1673–1685, Sep 2013.

[vK04] Joost van Kasteren. Retaining form in two dimensions: Depth - the missing dimension - is inside the human head, 2004.

[WGK10] Matthew Ward, Georges Grinstein, and Daniel Keim. *Human Perception and Information Processing in Interactive Data Visualization: Foundations, Techniques, and Application.* A K Peters, 2010.

[Wij12] Maarten W.A. Wijntjes. Probing pictorial relief: from experimental design to surface reconstruction. *Behavior Research Methods*, 44(1):135–143, 2012.

[WPG14] Susan G. Wardle, Stephen Palmisano, and Barbara J. Gillam. Monocular and binocular edges enhance the perception of stereoscopic slant. *Vision Research*, 100(0):113 – 123, 2014.

[ZLC94] G. Lee Zimmerman, Gordon E. Legge, and Patrick Cavanagh. Pictorial depth cues: a new slant. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 12(1):17–26, May 1994.

Glossary

accuracy Accuracy refers to the closeness of a measured value to a standard or known value, i.e. the ground truth. If for example in reality it is 30.0 C outside and a temperature sensor reads 30.1 C, then than sensor is relatively accurate. Accuracy is independent of precision. xi, 1, 12, 15, 21, 26, 29–31, 34, 37

albedo Also referred to as "reflection coefficient", is the diffuse reflectivity or reflecting power of a surface. 7

ANOVA Analysis of variance, a set of statistical models to analyze the differences between group means and their associated procedures, e.g. variation among and between groups. 29, 39

CGI Computer-generated imagery is the application of computer graphics to create images. 21

curvature Describes the amount by which a geometric object deviates from being flat. 14, 16

disparity Binocular disparity refers to difference in image location of an object seen by the left and right eyes. 7, 15, 17, 32, 34, 38

ground truth refers to the absolute truth of something, e.g. the actual surface normal. 1, 8, 12, 25

motion parallax Describes the apparent change in position of an object, if an observer changes his own position. 5, 7

precision Precision refers to the closeness of two or more measurements to each other. If for example you measure the temperature ten times and get 30.0 C each time, then the measurement is very precise. Precision is independent of accuracy. 21, 32, 33

retina A light-sensitive layer of tissue, lining the inner surface of the eye, which is responsible for transforming visual light stimuli into nerve impulses. 3, 4

shading The simulation of surface properties of objects in a virtual scene. 1

surface normal In the three-dimensional case a surface normal to a surface at a point P is a vector that is perpendicular to the tangent plane to that surface at P. 7