Untersuchungen zur visuellen Wahrnehmung der Form- und Materialeigenschaften dreidimensionaler transparenter Objekte

Dissertation
zur Erlangung des Doktorgrades
der Philosophischen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Nick Schlüter

Kiel, Februar 2019

Erstgutachter:	PD Dr. Franz Faul
Zweitgutachter:	PD Dr. Jürgen Golz
Drittgutachter:	Prof. Dr. Vebjørn Ekroll
Tag der mündlichen Prüfung:	2019-06-17
Durch den Prodekan für Studium und Lehre, Prof. Dr. Ulrich Müller, zum Druck genehmigt:	2019-06-19

Inhaltsverzeichnis

Kapitel 1

Begleittext

1.1 Einleitung

Im Alltag nehmen wir scheinbar mühelos die uns umgebenden Dinge und ihre Eigenschaften wahr: ihre Form, ihr Material, ihre Oberflächenbeschaffenheit, ihre Farbe, ihren Geruch, ihre Temperatur und so weiter. Darüber hinaus können wir zum Beispiel Lebewesen von unbelebten Dingen unterscheiden und haben insbesondere bei anderen Menschen oft eine Vorstellung ihrer Emotionen und Intentionen und können meist auf Anhieb einschätzen, ob sie uns feindlich oder wohlgesinnt sind. Allgemein ist unser subjektives Erleben von dem Eindruck geprägt, wir würden die Zustände und Vorgänge unserer Umwelt unmittelbar wahrnehmen können. Dabei scheinen unsere perzeptuellen Eindrücke quasi identisch mit den tatsächlichen Eigenschaften der distalen Außenwelt zu sein. Unsere Wahrnehmung scheint daher ein im Kern passiver Vorgang zu sein, der die Außenwelt auf eine wahrheitsgetreue Weise subjektiv erfahrbar macht. Unsere Alltagserfahrung gibt uns meist keinen Anlass, diese Einschätzung zu hinterfragen, sodass die menschliche Wahrnehmung zunächst kein sonderlich erklärungsbedürftiges Phänomen zu sein scheint.

Es gibt allerdings Fälle, in denen erkennbar wird, dass der Zusammenhang zwischen unseren subjektiven Wahrnehmungseindrücken und der objektiven Außenwelt komplexer ist, als intuitiv angenommen. Dies ist zum Beispiel beim sogenannten Chevreul-Stimulus (Chevreul, 1855) der Fall, der aus mehreren, nebeneinander angeordneten Balken aufsteigender Helligkeit besteht. In Abbildung 1.1a ist ein solcher Chevreul-Stimulus von einem inhomogenen Umfeld umgeben, dessen Helligkeit zusammen mit derjenigen der Balken nach rechts hin zunimmt (vgl. Geier & Hudák, 2011). Anders als die Helligkeit der Balken insgesamt, scheint die Helligkeit *innerhalb* der einzelnen Balken nach rechts hin leicht abzunehmen. Wird der Stimulus so verändert, dass der Helligkeitsgradient in der Umgebung nach links hin zunimmt, verändert sich die wahrgenommene Helligkeit der Balken (siehe Abbildung 1.1b). Zum einen scheint sie innerhalb der einzelnen Balken konstant zu sein, zum anderen scheinen sich die Helligkeiten der Balken insgesamt stärker voneinander zu unterscheiden. Diese Beob-

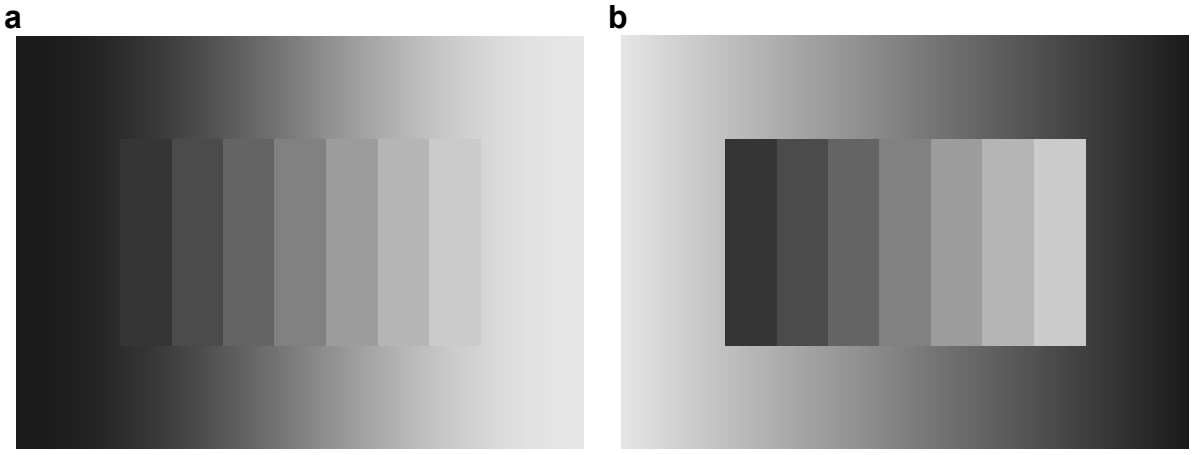

Abbildung 1.1: Am Beispiel des sogenannten Chevreul-Stimulus, der aus mehreren, nebeneinander angeordneten Balken aufsteigender Helligkeit besteht (Chevreul, 1855), lässt sich demonstrieren, dass die menschliche Wahrnehmung kein passiver und wahrheitsgetreuer Vorgang ist. **(a)** Ist der Chevreul-Stimulus von einem gleichläufigen Helligkeitsgradienten umgeben (vgl. Geier & Hudák, 2011), scheint die Helligkeit innerhalb der einzelnen Balken nach rechts hin leicht abzunehmen. **(b)** Wird der Gradient in der Umgebung an der Vertikalen gespiegelt, verändert sich die Erscheinung der Balken, obwohl ihre tatsächlichen Eigenschaften unverändert sind. Zum einen scheint die Helligkeit innerhalb jedes Balkens homogen zu sein, zum anderen scheinen sich die Helligkeiten der Balken insgesamt stärker voneinander zu unterscheiden.

achtung ist deshalb überraschend, weil die Balken selbst durch die Richtungsumkehr des sie umgebenden Helligkeitsgradienten nicht verändert wurden. Dass die Balken trotzdem in den beiden betrachteten Fällen inkonsistente Helligkeitseindrücke hervorrufen zeigt eindrücklich, „that perception is not just a simple registration of objective reality“ (Palmer, 1999, S. 9) und dass unsere Wahrnehmungseindrücke nicht unmittelbar die intrinsischen Eigenschaften der Außenwelt widerspiegeln. In der aktuellen Forschung wird Wahrnehmung daher meist als Folge eines konstruktiven Prozesses angesehen. Aus dieser Perspektive unterscheiden sich die subjektiven Eindrücke der Balken in den zwei soeben beschriebenen Situationen, weil diese auf unterschiedliche Weise konstruiert werden (für einen Erklärungsansatz siehe Schlüter und Golz, 2015). Darüber hinaus zeigt das betrachtete Beispiel, dass die menschliche Wahrnehmung nicht in dem Sinn wahrheitsgetreu ist, wie es aus subjektiver Sicht den Anschein macht: In Abbildung 1.1a ist die wahrgenommene Helligkeit innerhalb der einzelnen Balken inhomogen, obwohl die tatsächliche Helligkeit der Balken (das heißt ihre Leuchtdichte) konstant ist. Gleichwohl scheinen subjektive und objektive Eigenschaften in einer bestimmten Beziehung zueinander zu stehen: Die Zunahme der wahrgenommenen Helligkeit der Balken insgesamt korrespondiert mit der tatsächlichen Zunahme ihrer Leuchtdichte.

Die soeben angesprochene Helligkeitswahrnehmung ist nur ein Beispiel für Fälle, in denen sich unsere subjektiven Wahrnehmungseindrücke verhältnismäßig direkt auf physikalisch beschreibbare Eigenschaften der Außenwelt beziehen. Auch für das subjektive Attribut „Farbe“

besteht ein derartiger Zusammenhang. Bei isoliert dargebotenen Lichtern ist der subjektive Farbeindruck in eindeutiger Weise durch die spektrale Energieverteilung der Lichter bestimmt. Allerdings kann ein bestimmter Farbeindruck durch eine ganze Klasse von Lichtern mit unterschiedlichen Spektren hervorgerufen werden. Die subjektiven Farbeindrücke hängen daher zwar mit den physikalischen Spektren zusammen, stellen jedoch keinesfalls eine exakte Repräsentation dieser dar. Im Alltag gelangt Licht meist nicht direkt zum Auge, sondern wird von Oberflächen reflektiert. Das ins Auge gelangende Licht hängt hierbei nicht nur von dem Beleuchtungsspektrum der Lichtquelle, sondern auch vom Reflektanzspektrum der Oberfläche ab, das als deren „intrinsische Farbe“ bezeichnet werden könnte. Im subjektiven Farbeindruck spiegelt sich diese Dualität ebenfalls wieder: Wir sehen häufig nicht nur die Farbe der Oberfläche, sondern auch die Farbe der Beleuchtung (vgl. Mausfeld, 2003). Dass unsere Farbeindrücke nur mittelbar mit den physikalischen Eigenschaften der Außenwelt zusammenhängen, scheint im Alltag kein Problem darzustellen. Ein Grund dafür könnte sein, dass die natürlicherweise vorkommenden Spektren oft relativ glatte Verläufe mit breiten Energie- beziehungsweise Reflektanzmaxima besitzen, die eine gute Vorhersage der Interaktion von Licht und Reflektanz erlauben. Der Zusammenhang zwischen unseren Farbeindrücken und den physikalischen Größen Reflektanz und Beleuchtung ist zwar nicht exakt, aber doch so eng, dass eine erfolgreiche Interaktion mit der Außenwelt möglich ist.

Nicht alle unsere Wahrnehmungseindrücke beziehen sich auf physikalisch beschreibbare Eigenschaften der Außenwelt. Mitunter stehen sie mit viel abstrakteren Aspekten unserer Umwelt in Verbindung. Dies ist zum Beispiel der Fall, wenn wir die Emotionen und Intentionen anderer Menschen wahrnehmen. Der Wahrnehmungseindruck bezieht sich hier auf den hypothetischen mentalen Zustand eines anderen Lebewesens, der sich nach heutiger Vorstellung erst mittelbar aus komplexen biochemischen Vorgängen in dessen Körper ergibt. Ähnlich wie bei der Wahrnehmung von Farbe kann auch dieser Zusammenhang insofern als funktional bezeichnet werden, als er uns eine erfolgreiche Interaktion mit unserer Umwelt (und den in ihr lebenden anderen Menschen) ermöglicht.

Unabhängig von der konkreten Beschaffenheit des Zusammenhangs zwischen den subjektiven Wahrnehmungseindrücken und der objektiven Außenwelt deuten die bisherigen Beispiele darauf hin, dass zwischen beiden in der Regel eine funktionale Beziehung besteht. In diesem Sinn repräsentieren unsere Wahrnehmungseindrücke die für uns relevanten Aspekte der Außenwelt auf eine Weise, die es uns erlaubt, erfolgreich mit unserer Umwelt zu interagieren. Unsere Wahrnehmung kann daher als aktiver Prozess aufgefasst werden, der uns Kenntnisse über die Außenwelt liefert und uns auf diese Weise an unsere Umwelt koppelt.

1.2 Material- und Formwahrnehmung

Aus den bisherigen Betrachtungen geht hervor, dass die Wahrnehmung als Teil einer funktionalen Kopplung zwischen dem Organismus auf der einen und der Außenwelt auf der anderen Seite aufgefasst werden kann. Zwei für diese Kopplung entscheidende Aspekte unserer Umwelt, die im Mittelpunkt der vorliegenden Arbeit stehen, sind die Form und das Material von Objekten. Beide bestimmen wesentlich, wie wir mit der Welt interagieren können. Kenntnisse über die Form und das Material von Dingen zu besitzen scheint daher eine grundlegende Voraussetzung für eine erfolgreiche Interaktion mit der Außenwelt zu sein. Bereits kleinere Abweichungen zwischen wahrgenommenen und tatsächlichen Form- beziehungsweise Materialeigenschaften können negative Folgen haben. Dazu gehören zum Beispiel Kollisionen mit Gegenständen, deren räumliche Ausdehnung wir unterschätzt haben, Probleme beim Greifen nach Dingen, deren räumliche Ausdehnung wir überschätzt haben oder ungeeignete Interaktionen mit Dingen, deren Material wir falsch wahrgenommen haben. Pizlo (2008) schreibt insbesondere der Formwahrnehmung eine besondere Bedeutung zu. Die Form sei „the only perceptual property that has sufficient complexity to allow an object to be identified“ (Pizlo, 2008, S. 1) und unterscheide sich damit deutlich von allen anderen perzeptuellen Eigenschaften.

Unsere Alltagserfahrung zeigt sehr deutlich, dass Aspekte der Form und des Materials von Dingen eine wesentliche Rolle in unserem subjektiven Erleben der Welt spielen. Unter natürlichen Umständen haben wir meist einen reichhaltigen Wahrnehmungseindruck von Objekten, der sowohl ihre Form als auch ihre Materialeigenschaften einschließt. Wenngleich wir die Form eines Objekts zum Beispiel auch ertasten können, haben wir in der Regel bereits allein auf Basis unseres Sehsinns eine relativ genaue Vorstellung von ihr. Dies betrifft die räumliche Ausdehnung des Objekts als Ganzes ebenso wie lokale Variationen seiner Oberflächenausrichtung. Im Allgemeinen umfasst dieser Eindruck auch die nicht direkt sichtbaren Teile eines Objekts, wie etwa seine Rückseite oder anderweitig verdeckte Teile. Die Interpretation des zum Auge gelangenden Lichts im Sinn einer dreidimensionalen Form eines distalen Objekts scheint ein derart grundlegendes Merkmal unserer Wahrnehmung zu sein, dass selbst stark vereinfachte Situationen, wie zum Beispiel Strichzeichnungen, solche Eindrücke hervorrufen können (Cole et al., 2009; Mamassian & Kersten, 1996; Norman & Raines, 2002). Auch bezüglich der Materialeigenschaften der uns umgebenden Dinge haben wir meist eine sehr genaue Vorstellung. Wir können zum Beispiel wahrnehmen, ob die Oberfläche von Objekten rau oder glatt ist und ob das Material, aus dem sie bestehen, hart oder weich ist. Auf Basis solcher Wahrnehmungseindrücke können wir zum Beispiel Früchte von Steinen und Papier von Kunststoff unterscheiden (Sharan, Rosenholtz & Adelson, 2009, 2014). Ähnlich wie bei der Formwahrnehmung, entstehen Materialeindrücke oft bereits allein auf Basis unseres

Sehsinns, können sich jedoch durch die Integration von Informationen aus anderen Sinnen, zum Beispiel aus dem Tastsinn, weiter differenzieren (Bergmann Tiest & Kappers, 2007).

1.3 Mechanismen der Wahrnehmung

Eine aus wahrnehmungspsychologischer Sicht zentrale Frage ist, welche Prinzipien und Mechanismen der menschlichen Wahrnehmung zugrunde liegen. Diese Frage ist bereits seit vielen Jahrhunderten Teil sowohl der philosophischen als auch der naturwissenschaftlichen Auseinandersetzung mit dem menschlichen Geist. Entsprechend vielfältig sind die theoretischen Zugänge und Analyseebenen. Wenn die menschliche Wahrnehmung als aktiver Prozess aufgefasst wird, der uns an die Außenwelt koppelt, stellt sich die Frage, wie diese Kopplung zwischen subjektiver und objektiver Welt konkret funktioniert, das heißt auf welche Weise Wahrnehmungseindrücke generiert und Kenntnisse über die Außenwelt gewonnen werden.

Die vorliegende Arbeit beschäftigt sich mit dieser Frage im Hinblick auf die mit der Material- und Formwahrnehmung verbundenen visuellen Wahrnehmungsleistungen. Eine systematische Selbstbeobachtung ist dabei allerdings nicht ausreichend, da wir nur einen begrenzten introspektiven Zugang zu den mentalen Prozessen haben, die diesen Leistungen zugrunde liegen. Subjektiv stellt sich noch nicht einmal ein Problem, denn die Wahrnehmung scheint aus dieser Perspektive mühelos und unwillkürlich abzulaufen. Diese Abschottung ist mit der Annahme verträglich, dass es sich beim Wahrnehmungsmechanismus um ein „Modul“ handelt, das heißt ein eigenständiges funktionales System, dessen innere Vorgänge gegenüber anderen funktionalen Subsystemen, zum Beispiel dem Bewusstsein, abgeschirmt sind und nur solche Informationen ausgetauscht werden, die für die beteiligten Module notwendig und nutzbar sind (Fodor, 1983). In diesem Sinn ist lediglich das Endresultat des Wahrnehmungssystems Teil unseres subjektiven Erlebens.

1.4 Wahrnehmung als Informationsverarbeitung

Damit unsere Wahrnehmung überhaupt bestimmte Aspekte der Außenwelt abbilden kann, muss das Wahrnehmungssystem Informationen über die Außenwelt gewinnen können. Solche Informationen können in Form impliziter Annahmen oder Heuristiken über die Beschaffenheit der Welt im Wahrnehmungssystem repräsentiert sein. Beispielsweise wurde vorgeschlagen, das visuelle System gehe von der impliziten Annahme aus, dass Licht stets von oben komme (Ramachandran, 1988; Sun und Perona, 1998; siehe jedoch Morgenstern, Murray und Harris, 2011). Da solche impliziten Annahmen generelle Regularitäten betreffen, müssen sie natürlich nicht in jeder Situation korrekt sein.

Informationen über den tatsächlichen Zustand der Außenwelt in einer bestimmten Situation können ausschließlich von außen zum Organismus gelangen. Dies verdeutlicht, welche

entscheidende Rolle die sensorischen Eingangssignale für die Wahrnehmung spielen, schließlich sind sie die einzige Wirkung der Außenwelt auf den Organismus, die dieser erfassen kann. Die Eingangssignale ergeben sich aus der Erregung von Sinneszellen, die je nach Sinnesmodalität durch spezifische Reize geschieht. Im Fall der auditiven Wahrnehmung ist der spezifische Reiz eine Luftdruckschwankung, im Fall der visuellen Wahrnehmung das zum Auge gelangende Licht. Abgesehen von etwaigen impliziten Annahmen, müssen daher alle wesentlichen Informationen, auf deren Basis das Wahrnehmungssystem Kenntnisse über die Außenwelt generiert, bereits in den sensorischen Eingangssignalen enthalten sein. In diesem Sinn kann das Wahrnehmungssystem als informationsverarbeitendes System aufgefasst werden, das auf Basis der in den Eingangssignalen enthaltenen Informationen Wissen über die Außenwelt generiert und dieses teilweise in Form von subjektiv erlebbaren Wahrnehmungseindrücken zur Verfügung stellt. In Bezug auf das Sehen definiert Palmer (1999, S. 5) die Wahrnehmung entsprechend als „the process of acquiring knowledge about environmental objects and events by extracting information from the light they emit or reflect".

Um die Funktionsweise des Wahrnehmungssystems zu verstehen, erscheint es daher sinnvoll, zu analysieren, welche Informationen das sensorische Eingangssignal enthält. Ist bekannt, wie sich Informationen über einen bestimmten Aspekt der Außenwelt im Eingangsignal niederschlagen, so lässt sich ermitteln, welche Verarbeitungsschritte prinzipiell nötig wären, um Wissen über diesen Aspekt zu erlangen. Laut Marr (1982) spielen solche computationalen Betrachtungen eine entscheidende Rolle für das Verständnis eines informationsverarbeitenden Systems, denn „although algorithms and mechanisms are empirically more accessible, it is the top level, the level of computational theory, which is critically important from an information-processing point of view. The reason for this is that the nature of the computations that underlie perception depends more upon the computational problems that have to be solved than upon the particular hardware in which their solutions are implemented" (Marr, 1982, S. 27). Computationale Betrachtungen lassen sich wiederum in konkrete Hypothesen zur Funktionsweise des Wahrnehmungssystems überführen. Wurden bestimmte Charakteristika des Eingangssignals als Information identifiziert, die potentiell dazu genutzt werden kann, um einen von uns wahrnehmbaren Aspekt der Außenwelt zu erschließen, dann lässt sich empirisch untersuchen, inwiefern das Wahrnehmungssystem diese Charakteristika nutzt. Dazu wird zum Beispiel ihre Anwesenheit oder Beschaffenheit im Eingangssignal experimentell manipuliert. Im Unterschied zum klassischen psychophysikalischen Vorgehen, das den Fokus auf den Zusammenhang zwischen physikalischen Umgebungsbedingungen und subjektiven Empfindungen legt, liegt der Fokus dieser Forschungsstrategie explizit auf dem Informationsgehalt des Eingangssignals. Das skizzierte Vorgehen erlaubt daher das Aufstellen und Testen von Hypothesen, die sich explizit auf die Informationsverarbeitung des Wahrnehmungssystems beziehen.

Inwieweit die sensorischen Eingangssignale Informationen über bestimmte Aspekte der Außenwelt bereitstellen können, wird im Folgenden für die in dieser Arbeit behandelte visuelle Wahrnehmung beschrieben. In diesem Fall beruht das Eingangssignal auf dem zum Auge gelangenden Licht, das durch die Hornhaut, die Augenkammern, die Linse und den Glaskörper auf die Netzhaut projiziert wird und dort das sogenannte „retinale Bild“ erzeugt. Die Wirkung dieses retinalen Bilds (im Folgenden auch kurz „Bild“) auf die Photorezeptoren des Auges kann als das eigentliche Eingangssignal der Informationsverarbeitung durch das visuelle System aufgefasst werden. Zunächst erscheint es allerdings schwer vorstellbar, dass das vom Auge aufgenommene Licht so viel Information über die Außenwelt transportiert, dass auf seiner Basis unsere reichhaltigen Wahrnehmungseindrücke entstehen können. Im Folgenden wird daher zunächst die Entstehung des retinalen Bilds, die sogenannte Bildgenerierung, genauer beschrieben. Dieses Vorgehen ermöglicht es nachzuvollziehen, wann und wie sich Informationen über bestimmte Aspekte der Außenwelt im zum Auge gelangenden Lichtmuster niederschlagen. Die Analyse der Bildgenerierung wird zeigen, dass das retinale Bild zwar im Prinzip unzählige Informationen über die Außenwelt enthält, diese Informationen jedoch in der Regel nur implizit vorliegen und daher vom visuellen System aktiv extrahiert werden müssen.

1.5 Lichttransport im Fall opaker Materialien

Als Licht wird derjenige Teil des elektromagnetischen Spektrums bezeichnet, der die Sinneszellen der menschlichen Netzhaut erregen kann. Dies umfasst Photonen einer Frequenz von ungefähr 384 bis 789 Terahertz ($1\,\mathrm{THz} = 10^{12}\,\mathrm{Hz}$) beziehungsweise einer Wellenlänge von ungefähr 380 bis 780 Nanometern ($1\,\mathrm{nm} = 10^{-9}\,\mathrm{m}$; vgl. Stockman und Sharpe, 2000; Stockman, Sharpe und Fach, 1999). Der intrinsische Informationsgehalt eines einzelnen Photons beschränkt sich im Wesentlichen auf seinen Energiegehalt, der wiederum von seiner Frequenz abhängt. Darüber hinaus steckt in der bloßen Existenz eines Photons die Information, dass eine Lichtquelle existiert (oder zumindest existiert hat). Weitere Informationen ergeben sich vor allem aus den statistischen Eigenschaften einer größeren Anzahl von Photonen. Zum Beispiel hängt die relative Häufigkeit, mit der Photonen bestimmter Frequenzen in einem Licht vertreten sind, von der jeweiligen Lichtquelle ab. Die spektrale Energieverteilung eines Lichts enthält daher implizite Informationen über die Eigenschaften der dazugehörigen Lichtquelle. Auch die Ausbreitungsrichtung von Licht kann als Träger von Informationen aufgefasst werden: Unter der Annahme, dass Licht sich geradlinig ausbreitet, enthält seine momentane Ausbreitungsrichtung Informationen über seine, in entgegengesetzter Richtung liegende, Ursprungsrichtung. Da das zum Auge gelangende Licht systematisch auf die Netzhaut projiziert wird, hängt jede Stelle des retinalen Bilds mit einer bestimmten Ursprungsrichtung des Lichts relativ zur Orientierung des Auges zusammen. Von den Eigenschaften des retinalen

Bilds an einer bestimmten Stelle kann daher im Prinzip auf Eigenschaften der Außenwelt in der dazugehörigen Ursprungsrichtung geschlossen werden. Auf diese Weise ist es zum Beispiel möglich, eine Lichtquelle, deren Licht auf die Netzhaut trifft, räumlich zu verorten. Ohne Weiteres gilt dies allerdings nur für die Richtung, in der sie liegt, nicht jedoch für ihre Entfernung.

Im Allgemeinen gelangt Licht nicht ausschließlich auf direktem Wege von einer Lichtquelle zum Auge. Der Grund dafür ist, dass Licht mit den meisten Materialien in unserer Umgebung interagiert. Solche Interaktionen können sowohl an den Übergängen (Grenzflächen) zwischen unterschiedlichen Materialien auftreten, als auch innerhalb von Materialien. Sie können nicht nur die Ausbreitungsrichtung des Lichts ändern, sondern zum Beispiel auch seine spektrale Zusammensetzung. Diese Interaktionen mit Materialien bilden die Grundlage dafür, dass Licht implizite Informationen über unzählige Aspekte der Außenwelt transportiert.

Wenngleich der Fokus der vorliegenden Arbeit auf Materialien liegen wird, die von Licht durchdrungen werden können, beziehen sich die nachfolgenden Betrachtungen vorerst auf Situationen, in denen Licht nur mit der Oberfläche von Materialien interagiert. Da das Licht solche Materialien nicht durchdringt, werden diese auch als lichtundurchlässig, beziehungsweise opak bezeichnet. Am Beispiel solcher opaken Materialien lässt sich herleiten, wie Licht durch Eigenschaften unserer Umgebung beeinflusst wird und wie diese Einflüsse letztlich die Grundlage für den Informationsgehalt des retinalen Bilds bilden. Für die spätere Auseinandersetzung mit lichtdurchlässigen Materialien sind diese Betrachtungen insofern essentiell, als sich der Lichttransport im Fall opaker und lichtdurchlässiger Materialien zu großen Teilen auf dieselben physikalischen Phänomene zurückführen lässt.

Auf physikalischer Ebene lässt sich die Beeinflussung des Lichts durch Eigenschaften der Welt auf zwei grundlegende Phänomene zurückführen. Zum einen reflektieren viele opake Materialien einen Teil des auf sie treffenden Lichts. Diese Reflexionen führen zunächst dazu, dass sich die Ausbreitungsrichtung des Lichts ändert. Der relative Anteil des an der Oberfläche reflektierten Lichts hängt unter anderem vom Einfallswinkel des Lichts ab. In der Regel wird umso mehr Licht reflektiert, je flacher ein Lichtstrahl auf die Oberfläche trifft. Darüber hinaus bestimmt die Mikrostruktur der Oberfläche, in welche Richtung das einfallende Licht reflektiert wird (siehe Abbildung 1.2). Während glattere Oberflächen einfallendes Licht vornehmlich gerichtet in Spiegelungsrichtung reflektieren (spiegelnde Reflexion), wird Licht von rauen Oberflächen mit irregulärer Mikrostruktur weniger gerichtet in unterschiedliche Richtungen gestreut (diffuse Reflexion). Bei vielen Materialien hängt die Stärke der Reflexion, abgesehen vom Einfallswinkel, nicht von der Richtung ab, aus der das Licht auf ihre Oberfläche trifft (isotropes Reflexionsverhalten). Es existieren jedoch Materialien mit einem anisotropen Reflexionsverhalten, das durch richtungsabhängige Unterschiede in der Mikrostruktur ihrer Oberfläche hervorgerufen wird. Dazu gehören zum Beispiel Metalle mit gebürsteten Oberflächen, aber auch Textilien. Viele Materialien besitzen ein Reflexionsverhalten, das als

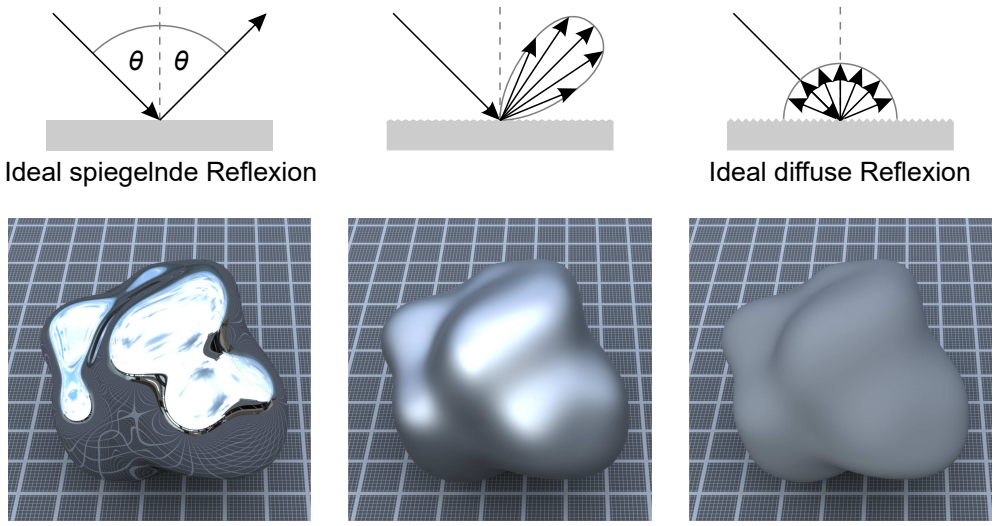

Abbildung 1.2: Einfluss der Mikrostruktur von Oberflächen auf ihr Reflexionsverhalten. Eine perfekt glatte Oberfläche (linkes Diagramm) würde einfallendes Licht ausschließlich exakt in Spiegelungsrichtung reflektieren (ideal spiegelnde Reflexion). Näherungsweise ist dies zum Beispiel bei polierten Metallen der Fall. Bei Oberflächen mit rauer Mikrostruktur wird einfallendes Licht hingegen um die Spiegelungsrichtung gestreut (mittleres Diagramm). Die Ausfallrichtung eines Teils des auftreffenden Lichts weicht dann mehr oder weniger stark von der eigentlichen Spiegelungsrichtung ab. Je rauer die Mikrostruktur der Oberfläche, desto ungerichteter erfolgt die Reflexion, sodass diese nicht mehr als spiegelnde, sondern als diffuse Reflexion bezeichnet wird. Im theoretischen Extremfall (rechtes Diagramm) wäre die Leuchtdichte des reflektierten Lichts in alle Richtungen konstant (ideal diffuse Reflexion; beschrieben durch das Lambertsche Kosinus-Gesetz). Näherungsweise ist dies etwa bei Papier oder Gips der Fall, deren Reflexionsverhalten zum Beispiel durch das Oren-Nayar-Modell beschrieben wird (Oren & Nayar, 1994). Die hier dargestellte Länge der reflektierten Lichtstrahlen und der Abstand ihrer Hüllkurve vom Reflexionspunkt entsprechen der Leuchtdichte in der jeweiligen Richtung.

Kombination von spiegelnder und diffuser Reflexion aufgefasst werden kann. Dazu zählen zum Beispiel Kunststoffe oder Materialien, deren Oberflächen lackiert sind.

Neben der Reflexion kann es an opaken Oberflächen zum Phänomen der Absorption kommen. Ein Teil des auftreffenden Lichts wird dabei vom Material aufgenommen und umgewandelt, zum Beispiel in thermische Energie. Je nach Eigenschaften des Materials werden Photonen unterschiedlicher Wellenlänge mit unterschiedlicher Wahrscheinlichkeit absorbiert, sodass sich nicht nur die Intensität, sondern auch die spektrale Verteilung des verbliebenen Lichts ändert (siehe Abbildung 1.3).

In der Regel kommt es an Oberflächen opaker Materialien sowohl zur Absorption als auch zur Reflexion, sodass sich sowohl die Ausbreitungsrichtung des auftreffenden Lichts ändert als auch seine spektrale Zusammensetzung. Wie viel des auftreffenden Lichts in Abhängigkeit von seinem Einfallswinkel und seiner Wellenlänge mit welcher Wahrscheinlichkeit absorbiert oder in bestimmte Richtungen reflektiert wird, wird durch sogenannte bidirektionale Reflektanz-

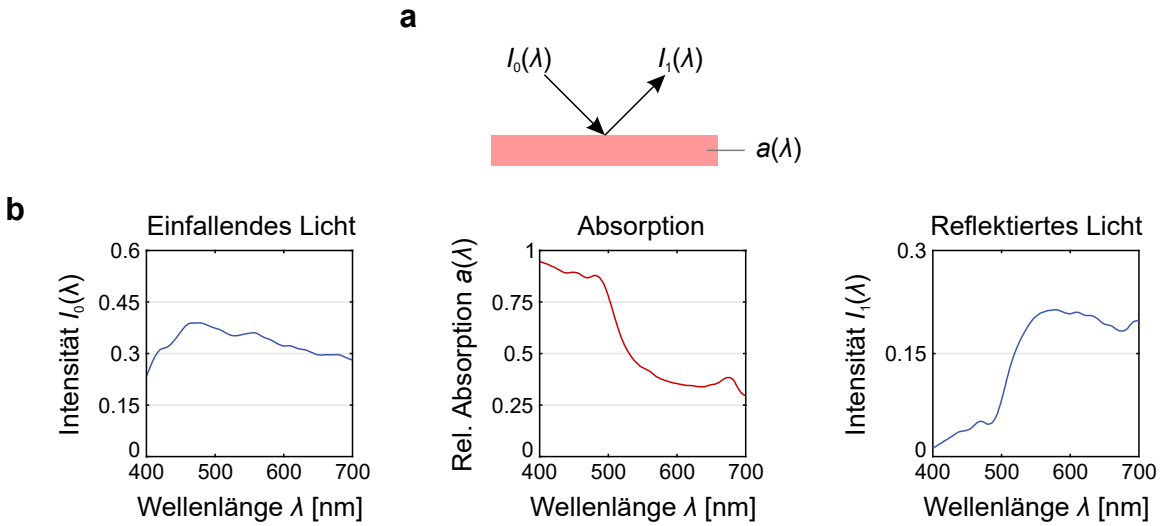

Abbildung 1.3: Absorption an opaken Oberflächen. **(a)** Trifft ein Lichtstrahl der Wellenlänge λ auf eine absorbierende Oberfläche, reduziert sich seine ursprüngliche Intensität $I_0(\lambda)$ in Abhängigkeit von den Absorptionseigenschaften der Oberfläche $a(\lambda)$ zu $I_1(\lambda) = I_0(\lambda) \times a(\lambda)$. **(b)** Sofern $a(\lambda)$ nicht für jede Wellenlänge λ denselben Wert hat, kann sich neben der Intensität auch die relative spektrale Verteilung des reflektierten Lichts ändern. Im hier gezeigten Beispiel wird Tageslicht (Taylor und Kerr, 1941, „sun plus sky“; linkes Diagramm) zum Teil von der Oberfläche eines Apfels absorbiert, der vor allem kurzwelliges Licht absorbiert (Vrhel, Gershon und Iwan, 1994, „Apple Yellow Delicious“; mittleres Diagramm). Im reflektierten Licht sind die mittel- und langwelligen Anteile entsprechend stärker vertreten (rechtes Diagramm).

verteilungsfunktionen (BRDF) beschrieben. Für unterschiedliche Materialien unterscheiden sich diese zum Teil erheblich.

Da der Einfluss, den Reflexion und Absorption auf das Licht haben, von den Eigenschaften des jeweiligen Materials abhängen, werden die Eigenschaften des Lichts nicht ausschließlich von den Eigenschaften der Lichtquelle bestimmt, von der es stammt, sondern auch von den Eigenschaften der Oberflächen, mit denen es interagiert hat. Da jede dieser Interaktionen einen systematischen Einfluss auf das Licht hat, enthält das retinale Bild im Prinzip Informationen über alle Oberflächen und Materialien, mit denen das Licht auf seinem Weg zum Auge interagiert hat. Ein prinzipielles Problem ergibt sich jedoch aus der Tatsache, dass das zum Auge gelangende Licht nur den Zustand *nach* allen Interaktionen abbildet, es jedoch keine expliziten Informationen darüber enthält, wann und wo es auf seinem bisherigen Weg welchen Einflüssen ausgesetzt war. Um auf Basis des retinalen Bilds konkrete Kenntnisse über die Außenwelt zu generieren, müssen sich die in ihm konfundierten Einflüsse daher zumindest teilweise voneinander trennen lassen.

1.5.1 Dominanz der letzten diffusen Reflexion

In der Regel wird die Konfundierung verschiedener Außenwelteinflüsse im zum Auge gelangenden Licht durch die Tatsache abgeschwächt, dass der Einfluss der jeweils letzten Oberfläche, mit der das Licht interagiert, besonders dominant ist. Der Grund dafür ist, dass viele natürliche Materialien Oberflächen mit teilweise diffuser Streuung besitzen und einfallendes Licht nicht ausschließlich in Spiegelungsrichtung reflektieren. Das von einer solchen Oberfläche in eine bestimmte Richtung reflektierte Licht stammt daher nicht nur aus der jeweiligen Spiegelungsrichtung, sondern aus mehr oder weniger stark darum streuenden Richtungen. Dieses aus unterschiedlichen Richtungen auftreffende Licht kann nicht nur von unterschiedlichen Lichtquellen mit unterschiedlichen Eigenschaften stammen, sondern darüber hinaus auf seinem bisherigen Weg unterschiedlichen Reflexions- und Absorptionseinflüssen durch andere Oberflächen ausgesetzt gewesen sein. Durch die Reflexion in eine gemeinsame Richtung mischen sich diese verschiedenen Lichter, sodass sich ihre bisherigen Interaktionen weniger systematisch in den Eigenschaften des reflektierten Lichts niederschlagen als ihre Interaktion mit der aktuellen Oberfläche. Die spektrale Energieverteilung eines Lichts wird daher überproportional stark durch die Absorptionseigenschaften derjenigen Oberfläche bestimmt, von der es zuletzt ungerichtet reflektiert wurde.

Die räumliche Konfiguration des insgesamt auf eine Stelle der Oberfläche treffenden Lichtmusters ändert sich durch die ungerichtete Reflexion in ein gänzlich anderes Muster des ausfallendes Lichts, sodass das reflektierte Licht ein anderes Bild transportiert als das auftreffende. Ungerichtete Oberflächenreflexionen führen daher zu einer Neuformierung des vom Licht transportierten Bilds. Abbildung 1.4 zeigt beispielhaft, wie stark sich die vom

Licht transportierten Bilder unmittelbar vor und nach einer diffusen Oberflächenreflexion unterscheiden.

Würden alle Oberflächen einfallendes Licht ausschließlich in Spiegelungsrichtung reflektieren, wäre der Einfluss aller Interaktionen des Lichts insofern gleichberechtigt, als es an keiner Oberfläche zu der soeben beschriebenen Mittelung der jeweils vorherigen Einflüsse käme. Abbildung 1.5 zeigt beispielhaft, wie sich die Erscheinung einer Szene ändert, wenn alle Oberflächen einfallendes Licht ausschließlich in Spiegelungsrichtung reflektieren. In diesem Fall werden die spektralen Eigenschaften des von den verschiedenen Oberflächen zum Auge gelangenden Lichts nicht mehr durch die Absorptionseigenschaften dieser Oberflächen dominiert, sondern sie ergeben sich aus einem gleichwertigen Einfluss der Absorptionseigenschaften aller Oberflächen, mit denen das Licht interagierte. Darüber hinaus fällt die räumliche Struktur des Bilds komplexer aus. Der Grund dafür ist, dass das Bild nicht nur die Konturen der Gegenstände abbildet, von deren Oberflächen das Licht unmittelbar (das heißt ohne weitere Reflexion) zum Auge gelangt, sondern auch Konturen der Gegenstände, die sich in den jeweiligen Oberflächen spiegeln. Diese gespiegelten Gegenstände zeigen auf ihren Oberflächen wiederum Spiegelbilder der Konturen anderer Oberflächen und so fort.

1.6 Informationen im retinalen Bild opaker Materialien

Auf Basis der beiden aus physikalischer Sicht zentralen Phänomene der Reflexion und Absorption kann das zum Auge gelangende Licht im Prinzip Informationen über alle Aspekte der Außenwelt beinhalten, die einen irgendwie gearteten Einfluss auf das Licht ausüben. Dies umfasst insbesondere auch eine im Prinzip unbegrenzte Menge mehr oder weniger abstrakter Aspekte, die nur indirekt auf den physikalischen Lichttransport wirken. Dazu gehören zum Beispiel auch die bereits angesprochenen emotionalen Zustände anderer Menschen. Ihr Einfluss auf den Lichttransport kann sich etwa aus einer Änderung der räumlichen Konfiguration der Gesichtsoberfläche aufgrund der Kontraktion bestimmter Muskeln ergeben, die wiederum systematisch mit bestimmten Emotionen zusammenhängen. Im retinalen Bild können sich alle derartigen Einflüsse als spezifische Regularitäten niederschlagen. Bestimmte Aspekte des retinalen Bilds hängen dann mehr oder weniger systematisch mit bestimmten Aspekten der Außenwelt zusammen. Im Prinzip können solche Regularitäten im retinalen Bild dem visuellen System daher als Hinweis („Cue“) auf diese Aspekte dienen. Wie die jeweiligen Regularitäten im retinalen Bild genau beschaffen sind und auf welche Weise sie mit Aspekten der Außenwelt zusammenhängen, unterscheidet sich zum Teil erheblich. Darüber hinaus ist der Zusammenhang in der Regel nicht eindeutig. Eine spezifische Regularität kann zum Beispiel durch verschiedene Außenweltaspekte beeinflusst sein, sodass von bestimmten Eigenschaften des retinalen Bilds nicht eindeutig auf einen spezifischen Aspekt geschlossen werden kann. Auch können sich manche Aspekte der Außenwelt in mehreren verschiedenen Regularitäten

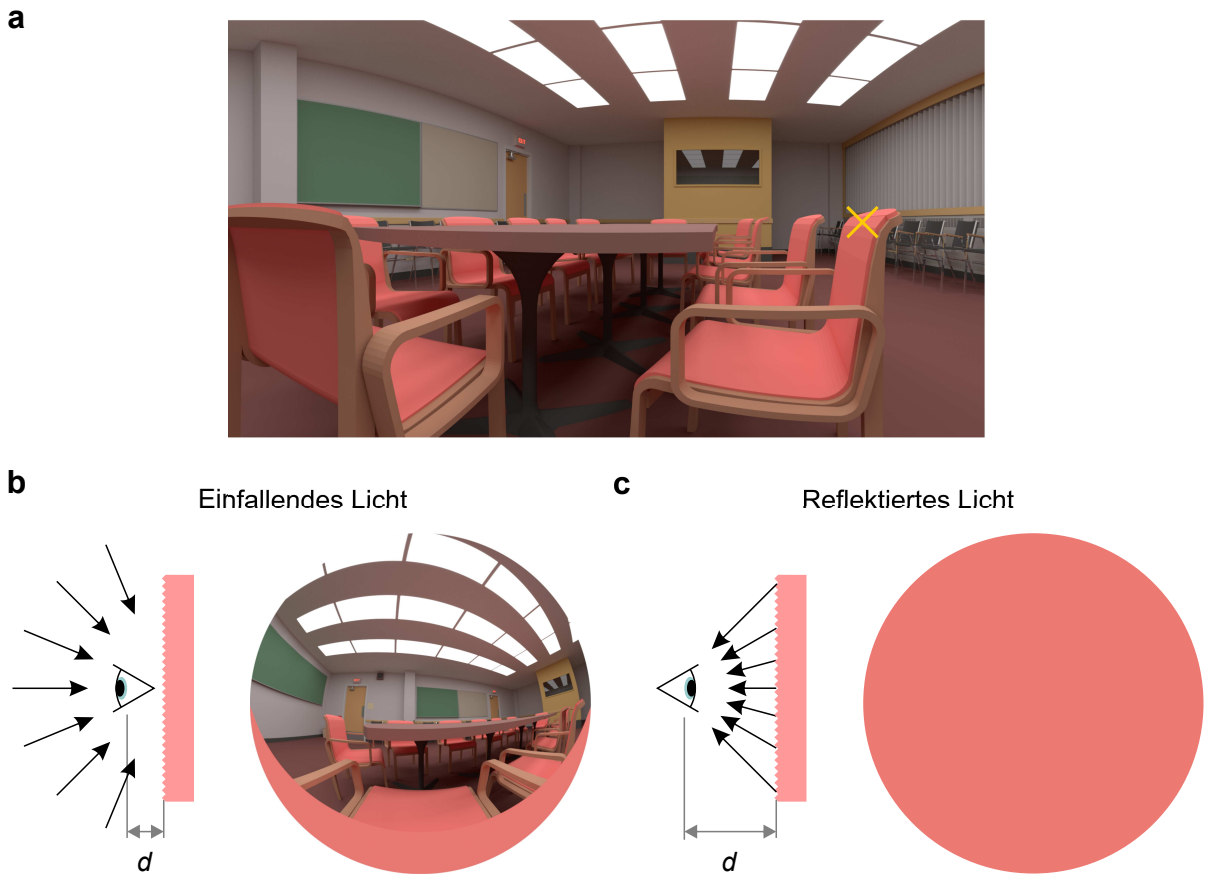

Abbildung 1.4: Illustration der Neuformierung des vom Licht transportierten Bilds durch diffuse Oberflächenreflexion. **(a)** Als Reflexionspunkt wird beispielhaft die im Bild durch ein gelbes Kreuz hervorgehobene Stelle auf der Rückenlehne eines Stuhls angenommen. Alle Bilder sind computergeneriert (basierend auf einem digitalen Modell von Grynberg und Ward, 1990; konvertiert für die Bildsynthese in Cycles, Blender Foundation, 2015, von Walter, 2014). **(b)** Unmittelbar vor der Reflexion (d minimal größer Null) transportiert das Licht ein Bild, dessen Eigenschaften von den Lichtquellen sowie den Reflexions- und Absorptionseinflüssen bestimmt werden, denen das Licht auf seinem bisherigen Weg ausgesetzt war. Die räumliche Struktur des Bilds ergibt sich aus den richtungsabhängigen Unterschieden der einfallenden Lichtstrahlen. **(c)** Durch die diffuse Reflexion mischen sich die aus unterschiedlichen Richtungen stammenden Lichtstrahlen und damit auch ihre Eigenschaften. Das unmittelbar nach der Reflexion (d minimal größer Null) vom Licht transportierte Bild wird daher überproportional stark durch die Absorptionseigenschaften der Oberfläche bestimmt, von der es reflektiert wurde (im hier gezeigten Beispiel absorbiert die Oberfläche vor allem kurzwelliges Licht).

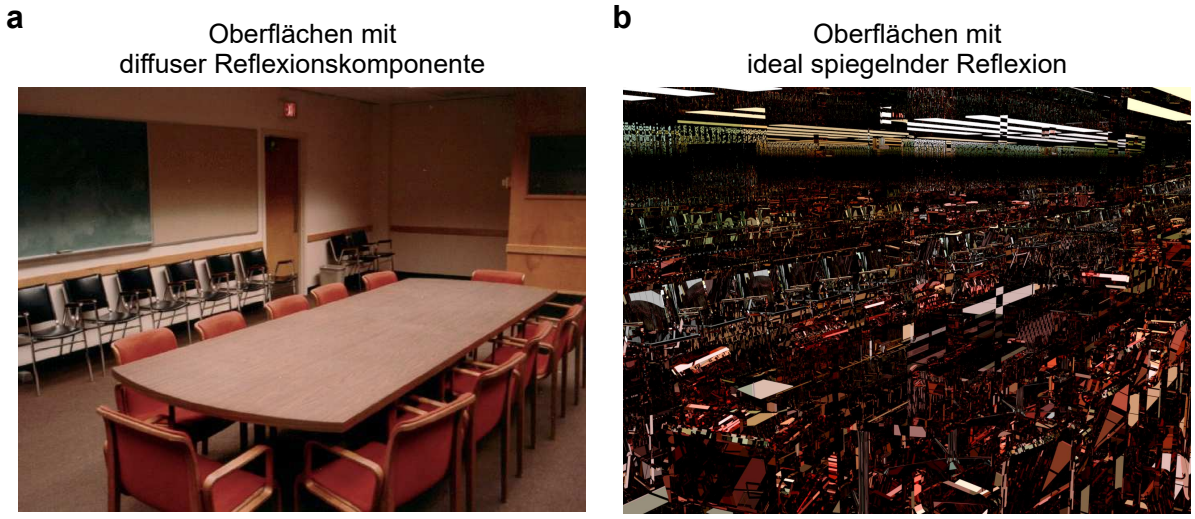

Abbildung 1.5: Einfluss der Art der Oberflächenreflexion auf die Struktur des vom Licht transportierten Bilds. **(a)** Sofern die meisten Oberflächen einer Szene eine diffuse Oberflächenreflexion besitzen, ergibt sich die räumliche Struktur des Bilds vor allem aus den Absorptionseigenschaften der Oberflächen, von denen das Licht *zuletzt* reflektiert wurde. Gezeigt ist die Fotografie eines Konferenzraums (Larson & Shakespeare, 2003, Tafel 2). **(b)** Wenn alle Oberflächen einer Szene einfallendes Licht ausschließlich in Spiegelungsrichtung reflektieren würden, ergäbe sich die räumliche Struktur des Bilds aus den Absorptionseigenschaften *aller* Oberflächen, von denen das Licht auf seinem Weg reflektiert wurde. Das Bild würde nicht nur die Konturen von direkt (das heißt ohne weitere Reflexion) projizierten Gegenständen beinhalten, sondern auch Konturen von Gegenständen, deren Licht ein oder mehrmals spiegelnd reflektiert wurde. Das gezeigte Bild ist computergeneriert und zeigt den in Abbildung 1.5a abgebildeten Konferenzraum aus vergleichbarer Perspektive (basierend auf einem digitalen Modell von Grynberg und Ward, 1990; konvertiert für die Bildsynthese in Cycles, Blender Foundation, 2015, von Walter, 2014).

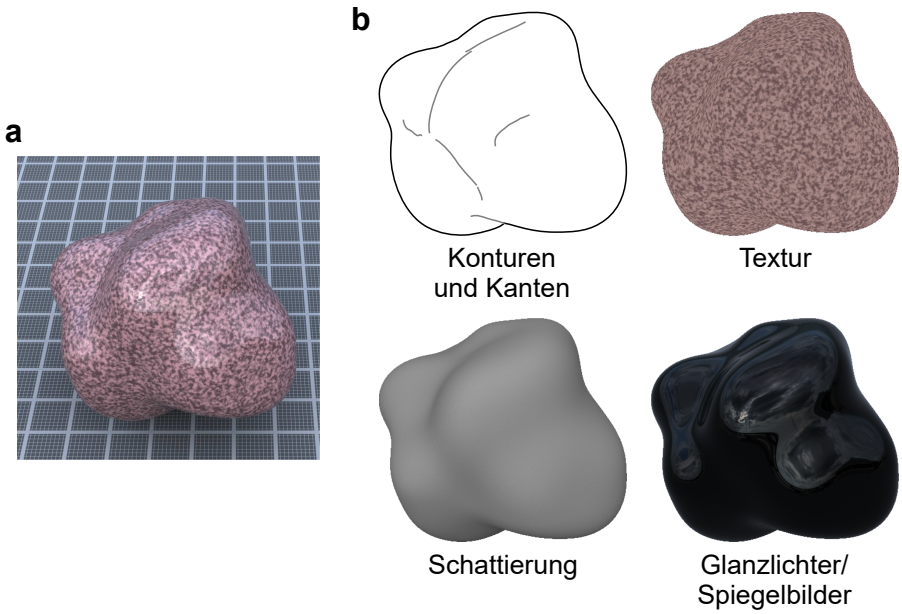

Abbildung 1.6: Übersicht über einige der bekannten Hinweise auf die Form opaker Objekte. **(a)** Beispiel eines irregulär geformten, texturierten Objekts mit diffusen und spiegelnden Reflexionen, wie es in ähnlicher Weise vielfach in Arbeiten zur Formwahrnehmung verwendet wird (zum Beispiel Norman, Todd & Phillips, 1995; Todd, Norman, Koenderink & Kappers, 1997). **(b)** Potentiell als Formhinweis nutzbare Bildregularitäten beziehen sich zum Beispiel auf die Kontur eines Objekts (oben links), die Dichte und Form seiner Texturelemente (oben rechts), seine durch diffuse Reflexionen verursachten Schattierungen (unten links) und durch spiegelnde Reflexionen verursachte Glanzlichter beziehungsweise Spiegelbilder (unten rechts).

niederschlagen. Von diesen müssen wiederum nicht alle in jeder Situation gleichermaßen gut geeignet sein, um bestimmte Kenntnisse über die Außenwelt zu erlangen.

Welche Regularitäten im retinalen Bild mit der Form von Objekten in Zusammenhang stehen und wie diese vom visuellen System verwendet werden, wurde bereits in einer Vielzahl von theoretischen und empirischen Arbeiten genauer untersucht. Wesentliche Hinweise auf die Form von Oberflächen ergeben sich unter anderem aus Konturen und Kanten, ihrer Textur, Schattierungen sowie aus Spiegelungen und Glanzlichtern (siehe Abbildung 1.6). Für die spätere Auseinandersetzung mit der Wahrnehmung lichtdurchlässiger Materialien sind diese Formhinweise und die entsprechenden Arbeiten insofern relevant, als sich einige der Hinweise auf identische oder ähnliche Weise auch im lichtdurchlässigen Fall identifizieren lassen. Es erscheint daher plausibel, dass zumindest einige der aus dem opaken Fall bekannten Mechanismen auch im lichtdurchlässigen Fall Anwendung finden.

1.6.1 Formhinweise aus Konturen und Kanten

Frühe Arbeiten zur Formwahrnehmung bezogen sich meist auf Bildkriterien, die sich allein aus der Geometrie eines Objekts ergeben. Dazu zählt insbesondere seine äußere Kontur im

Bild (siehe Abbildung 1.6b oben links, schwarze Linie). Neben der Analyse, welche Abschnitte äußerer Konturen besonders informativ sind (Attneave, 1954; Feldman & Singh, 2005; Kennedy & Domander, 1985; Norman, Phillips & Ross, 2001; Panis, De Winter, Vandekerckhove & Wagemans, 2008) entstanden eine Reihe computationaler Überlegungen, wie die Kontur als Formhinweis verwendet werden könnte (Barrow & Tenenbaum, 1981; Koenderink, 1984; Li, Pizlo & Steinman, 2009; Malik, 1987; Mamassian & Landy, 1998; Marr, 1977; Richards, Koenderink & Hoffman, 1987; Tse, 2002). Koenderink (1984) zeigte beispielsweise, dass konvexe und konkave Bereiche der Kontur oft auf konvexe beziehungsweise sattelförmige Oberflächenbereiche hinweisen. Neben der äußeren Kontur wurden auch Kanten innerhalb des Objektbereichs, die sich entweder durch Tiefensprünge der Oberfläche (das heißt durch Selbstverdeckung) oder abrupte Änderungen der Oberflächenausrichtung ergeben (siehe Abbildung 1.6b oben links, graue Linien), als relevant für die Formwahrnehmung vorgeschlagen (Barrow & Tenenbaum, 1981; Waltz, 1975). Mehrere Arbeiten konnten nicht nur zeigen, dass Konturen und Kanten oft ausreichen, um einen dreidimensionalen Formeindruck zu erzeugen (Cole et al., 2009; Mamassian & Kersten, 1996; Norman & Raines, 2002), sondern auch, dass sich dieser durch andere Formhinweise wie Schattierung nur noch wenig verbessert (Christou & Koenderink, 1997; Koenderink, van Doorn, Christou & Lappin, 1996).

1.6.2 Formhinweise aus Textur

Eine weitere Informationsquelle für die Form von opaken Objekten ist ihre projizierte Textur im Bild. Diese Textur kann sich sowohl aus den geometrischen Eigenschaften des Objekts ergeben (wie im Fall eines Golfballs mit seinen zahlreichen Vertiefungen) als auch aus der räumlichen Variation seiner spektralen Reflektanzeigenschaften, das heißt aus seiner intrinsischen Textur (siehe Abbildung 1.6b oben rechts). In einer Vielzahl von Arbeiten wurde der reguläre Zusammenhang zwischen bestimmten Eigenschaften dieser projizierten Texturen und der Form sowie ihre Nutzung als Formhinweis untersucht (Fleming, Holtmann-Rice & Bülthoff, 2011a; Gårding, 1990; Gibson, 1950; Malik & Rosenholtz, 1997; Todd & Akerstrom, 1987). Einige der Ansätze beziehen sich dabei stärker auf globale statistische Regularitäten der Textur, wie die Dichte von Texturelementen (zum Beispiel Blake, Bülthoff & Sheinberg, 1993; Cutting & Millard, 1984; Gibson, 1950; Rosenholtz & Malik, 1997; Stevens, 1981; Todd, Oomes, Koenderink & Kappers, 2004; Witkin, 1981), während andere sich auf lokale Eigenschaften einzelner Texturelemente, wie ihre Größe und Form, beziehen (zum Beispiel Lobay & Forsyth, 2006; Stevens, 1980). Allen Ansätzen ist gemein, dass sie bestimmte Annahmen bezüglich der intrinsischen Textur machen, zum Beispiel in Bezug auf die gleichmäßige Verteilung der Texturelemente oder ihre konstante mittlere Größe. Bezüglich der Generalisierbarkeit der Resultate dieser Arbeiten gilt es zu bedenken, dass sie sich oft nur auf flache (Stevens, 1981) oder einfach gekrümmte Oberflächen (Gårding, 1992) bezogen haben. Erst die Berücksichtigung

komplexerer Objekte mit mehrfach gekrümmten Oberflächen (zum Beispiel Todd et al., 2004; Todd & Thaler, 2010) erlaubte hinreichend allgemeine Aussagen zur Rolle der Textur für die Formwahrnehmung.

1.6.3 Formhinweise aus Schattierung

Weitere Formhinweise ergeben sich aus der Mikrostruktur einer Oberfläche und der Art wie diese Licht reflektiert. Wie bereits beschrieben, reflektieren raue Oberflächen mit unregelmäßiger Mikrostruktur einfallendes Licht eher diffus. Wie viel Licht von solchen Oberflächen reflektiert wird, hängt vor allem von ihrer Ausrichtung zur Lichtquelle ab. Stärker zu einer Lichtquelle hin ausgerichtete Oberflächen werden stärker beleuchtet und reflektieren mehr Licht als Oberflächen, die von der Lichtquelle weggeneigt sind. Liegt eine räumliche Variation der Ausrichtung einer Oberfläche zur Lichtquelle vor, so erzeugen die dadurch hervorgerufenen Variationen der reflektierten Lichtmenge im Bild nahezu ortsfeste Schattierungen, die hauptsächlich von der relativen Position der Lichtquelle und kaum von der Betrachtungsposition abhängen (siehe Abbildung 1.6b unten links). Aufgrund des Zusammenhangs dieser Schattierungen mit der Oberflächenausrichtung wurden beginnend mit Horn (1970) zahlreiche computationale Ansätze vorgestellt, wie Schattierungen aufgrund von diffusen Oberflächenreflexionen zur Formerkennung verwendet werden können (Horn, 1975, 1990; Ikeuchi & Horn, 1981; Koenderink & van Doorn, 1980; Lee & Rosenfeld, 1985; Pentland, 1989). Verglichen mit den Formhinweisen aus Konturen, Kanten und Texturen ist die Nutzung von Schattierungen als Formhinweis insofern computational komplexer, als diese maßgeblich von Eigenschaften der Beleuchtung beeinflusst werden. Ändern sich die Anzahl der Lichtquellen und/oder ihre räumlichen Positionen und räumlichen Ausdehnungen, ändern sich auch die Schattierungen. Viele Arbeiten zu Schattierungen als Formhinweis bezogen sich daher oft zusätzlich auf mögliche Annahmen des visuellen Systems bezüglich der Beleuchtung (Gerardin, de Montalembert & Mamassian, 2007; Kleffner & Ramachandran, 1992; Mingolla & Todd, 1986; O'Shea, Banks & Agrawala, 2008; Ramachandran, 1988; van Doorn, Koenderink & Wagemans, 2011).

1.6.4 Formhinweise aus Spiegelbildern und Glanzlichtern

Auch Oberflächen, die einfallendes Licht hauptsächlich in Spiegelungsrichtung reflektieren, können im Bild Regularitäten erzeugen, die zur Formwahrnehmung genutzt werden können. Sofern eine Oberfläche einfallendes Licht exakt in Spiegelungsrichtung reflektiert (ideal spiegelnde Reflexion), treten auf ihr Spiegelbilder der Umgebung auf (siehe Abbildung 1.6b unten rechts). Oft erscheinen diese Spiegelbilder mehr oder weniger stark verzerrt. *Was* sich auf dem Objekt spiegelt hängt insbesondere von seiner Umgebung, einschließlich der Beleuchtung, und der Position des Beobachters ab. *Wie* sich die Umgebung spiegelt wird hingegen maßgeblich von der Form des Objekts beeinflusst. Beispielsweise erscheint an

Stellen mit starker Oberflächenkrümmung das Spiegelbild stärker verzerrt als an Stellen mit schwacher oder keiner Krümmung. Wie genau spiegelnde Reflexionen mit der Form in Verbindung stehen und wie sie zur Formerkennung verwendet werden können, wurde bereits ausführlicher untersucht (Oren & Nayar, 1997; Savarese, Chen & Perona, 2004a, 2005; Savarese, Fei-Fei & Perona, 2004b; Savarese & Perona, 2001, 2002). Fleming, Torralba und Adelson (2004) sowie Muryy, Welchman, Blake und Fleming (2013) beziehen sich darüber hinaus explizit auf die Frage, wie das visuelle System Spiegelungen als Formhinweis verwendet und auf welche Regularitäten der Spiegelungen es sich dabei bezieht. Sie betonen dabei den Zusammenhang, der zwischen der Richtungsabhängigkeit der Verzerrungen im Bild und den Krümmungen der betrachteten Oberfläche besteht. Bewegen sich Objekt und/oder Betrachter, ergeben sich daraus weitere Regularitäten im Bild, die zur Formschätzung verwendet werden können. Diese beziehen sich zum Beispiel auf den optischen Fluss, mit der sich die von der spiegelnden Oberfläche abgebildete Umgebung im Bild bewegt (Adato, Vasilyev, Ben-Shahar & Zickler, 2007; Adato, Vasilyev, Zickler & Ben-Shahar, 2010; Dövencioglu, Ben-Shahar, Barla & Doerschner, 2017).

Streut das reflektierte Licht stärker um die Spiegelungsrichtung, werden die Spiegelbilder zunehmend diffus. Objekte aus solchen Materialien erscheinen oft matt glänzend und können auf ihrer Oberfläche mehr oder weniger stark ausgeprägte Glanzlichter besitzen. Auch diese diffuseren Glanzlichter können als Formhinweis dienen, da ihre Anzahl, Position, Form und Größe mit der Objektform in Zusammenhang steht. Wie die Spiegelungen aus ideal spiegelnder Reflexion unterliegen auch sie dem Einfluss durch die Beleuchtung und die Position des Beobachters. Beispielsweise begünstigen räumlich begrenzte Lichtquellen das Auftreten einzelner Glanzlichter, homogene Beleuchtungen hingegen nicht. Oft wurde die Rolle von Glanzlichtern daher an isolierten Objekten und Oberflächen unter punktförmiger oder gerichteter Beleuchtung untersucht (Blake & Bülthoff, 1990, 1991; Mingolla & Todd, 1986; Nefs, 2008; Norman, Todd & Orban, 2004b; Norman, Todd & Phillips, 1995).

Besitzt die Oberfläche ein Reflexionsverhalten, das als Kombination von spiegelnder und diffuser Reflexion aufgefasst werden kann, ergeben sich im Bild zwei Komponenten: Schattierungen aufgrund diffuser Reflexion und Spiegelungen oder Glanzlichter aufgrund (ideal) spiegelnder Reflexion. Durch Variation der Parameter solcher Materialien wurde zum Beispiel der relative Beitrag von Spiegelungen und Glanzlichtern zur Formwahrnehmung untersucht (vgl. Mingolla & Todd, 1986; Nefs, 2008; Norman et al., 1995).

1.7 Lichttransport im Fall lichtdurchlässiger Materialien

Die bisherigen Betrachtungen haben gezeigt, welche Regularitäten im retinalen Bild mit der Form in Verbindung stehen und welche davon vom visuellen System als Hinweis auf die Form verwendet werden. Außerdem hat die vorhergehende Betrachtung zum Lichttransport gezeigt,

dass diese und andere Leistungen des visuellen Systems offenbar ganz wesentlich auf der Tatsache beruhen, dass Licht an den Oberflächen vieler Materialien in unserer Umgebung reflektiert und absorbiert wird. Im Allgemeinen interagiert Licht jedoch nicht nur mit der unmittelbaren Oberfläche von Materialien. In vielen Fällen dringt Licht mehr oder weniger tief in ein Material ein oder durchquert dieses sogar vollständig. Im Folgenden wird erörtert, inwieweit sich der Lichttransport bei lichtdurchlässigen Materialien von dem bei opaken Materialien unterscheidet.

Eine wichtige und folgenreiche Eigenschaft lichtdurchlässiger Materialien ist, dass sich die Ausbreitungsrichtung des Lichts ändert, wenn es ihre Grenzfläche passiert (siehe Abbildungen 1.7a und 1.7b). Die Stärke dieser sogenannten Lichtbrechung beziehungsweise Refraktion hängt vom Einfallswinkel des Lichts ab und von der optischen Dichte des Materials und der seiner Umgebung (siehe Abbildung 1.7c). Die optische Dichte bestimmt, um welchen Faktor die Phasengeschwindigkeit des sich in dem Material ausbreitenden Lichts kleiner ist als im Vakuum und wird durch den sogenannten Brechungsindex (R) angegeben. Lichtdurchlässige Materialien, wie zum Beispiel Wasser oder Glas, aber auch Luft, besitzen jeweils einen spezifischen Brechungsindex. Der Brechungsindex ist also ein Materialkonstante. Für Luft ist der Brechungsindex zum Beispiel nur minimal größer als 1, für Wasser 1.33 und für Diamant 2.42. Die optische Dichte eines Materials ist zwar oft, jedoch keineswegs immer proportional zu seiner stofflichen Dichte (Bergmann, Schaefer & Bergmann, 1993, S. 50).

Darüber hinaus ist die Absorption bei lichtdurchlässigen Materialien in der Regel deutlich schwächer ausgeprägt als im opaken Fall, sodass das auftreffende Licht nicht bereits nahe der Oberfläche vollständig absorbiert wird, sondern nur teilweise innerhalb des Materials. Wie stark sich durch diese Volumenabsorption die Intensität des Lichts und seine spektrale Verteilung ändern, hängt neben den Absorptionseigenschaften des Materials auch von der Länge der Strecke ab, die das Licht innerhalb des Materials zurücklegt (siehe Abbildung 1.8).

Ähnlich wie opake Materialien reflektieren auch lichtdurchlässige Materialien einen Teil des auf sie treffenden Lichts. Wie groß dieser Anteil ist, hängt wie im opaken Fall vom Einfallswinkel des Lichts und der optischen Dichte des Materials selbst und der seiner Umgebung ab (siehe Abbildung 1.9a). Aufgrund ihrer Lichtdurchlässigkeit kann Licht nicht nur von außen auf eine lichtdurchlässige Oberfläche treffen und reflektiert werden, sondern auch von innen (siehe Abbildung 1.9b). Diese internen Reflexionen unterscheiden sich von den externen zum einen dadurch, dass sie eine andere Winkelabhängigkeit besitzen, zum anderen kann es bei den meisten von Luft umgebenen Materialien bei größeren Einfallswinkeln zur Totalreflexion kommen, bei der das auftreffende Licht vollständig reflektiert wird und das Material an der Reflexionsstelle nicht verlässt. Im Bild sind das intern und extern spiegelnd reflektierte Licht additiv überlagert (siehe Abbildung 1.9c).

Während die Mikrostruktur einer Oberfläche im opaken Fall vor allem bestimmt, wie stark das reflektierte Licht um die jeweilige Spiegelungsrichtung streut, bestimmt sie bei lichtdurch-

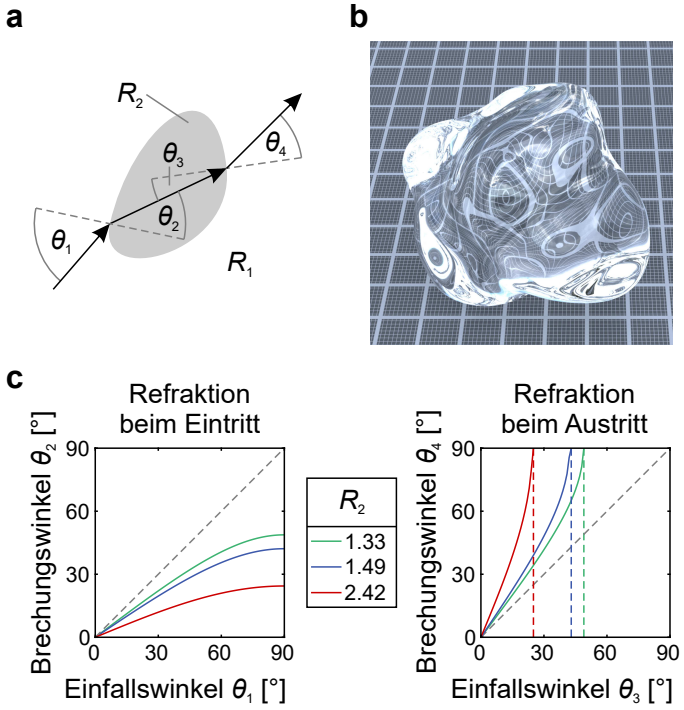

Abbildung 1.7: Lichtbrechung an lichtdurchlässigen Oberflächen. **(a)** Unterscheidet sich der Brechungsindex eines Materials (R_2) von dem seiner Umgebung (R_1), kommt es an seiner Oberfläche zur Lichtbrechung. Im hier gezeigten Beispiel ist $R_2 > R_1$, wie bei den meisten von Luft umgebenen Materialien. Trifft ein Lichtstrahl mit dem Winkel θ_1 von außen auf das Material, wird er in Richtung der Oberflächennormalen (gestrichelte, graue Linie) gebrochen. Der Brechungswinkel θ_2 ist dann kleiner als der Einfallswinkel θ_1 und beträgt $\theta_2 = \arcsin((R_1/R_2) \times \sin\theta_1)$ (Snelliussches Brechungsgesetz). Tritt das Licht wieder aus dem Material aus, wird es von der Oberflächennormalen weg gebrochen. Der Brechungswinkel θ_4 ist dann größer als der Einfallswinkel θ_3 und beträgt $\theta_4 = \arcsin((R_2/R_1) \times \sin\theta_3)$. In der Regel nimmt der Brechungsindex mit zunehmender Wellenlänge des Lichts ab, sodass kurzwelliges Licht stärker gebrochen wird als langwelliges. Dieser als Dispersion bezeichnete Effekt ist unter anderem für die chromatische Aberration von optischen Linsen verantwortlich, wird hier jedoch der Einfachheit halber vernachlässigt. Ebenso vernachlässigt werden Effekte, die sich aus der Polarisation des Lichts (das heißt der Schwingungsrichtungen seiner Photonen) ergeben. **(b)** Die Lichtbrechung moduliert die räumliche Konfiguration des durch das Objekt transmittierten Lichts und sorgt so für optische Verzerrungen des durch das Objekt sichtbaren Hintergrunds. **(c)** Sowohl beim Eintritt in ein lichtdurchlässiges Material als auch beim Austritt aus ihm ist die Stärke der Lichtbrechung umso größer, je größer die Differenz der beiden Brechungsindizes (R_2 und R_1) und je größer der Einfallswinkel des Lichts (θ_1 beziehungsweise θ_3). Beispielhaft sind hier die Einfalls- und Brechungswinkel für den Eintritt in (links) und den Austritt aus (rechts) unterschiedlichen von Luft ($R_1 \approx 1$) umgebenen Materialien dargestellt. An der Grenzfläche zum optisch dünneren Material kann es zum Phänomen der Totalreflexion kommen, wenn der Einfallswinkel einen bestimmten kritischen Wert ($\theta_c = \arcsin(R_1/R_2)$; senkrechte, gestrichelte Linien) überschreitet (Diamant: $\theta_c = 24.4°$; Glas: $\theta_c = 42.2°$; Wasser: $\theta_c = 48.8°$). Das Licht wird dann vollständig an der Grenzfläche reflektiert und der nach außen transmittierte Anteil sinkt auf Null.

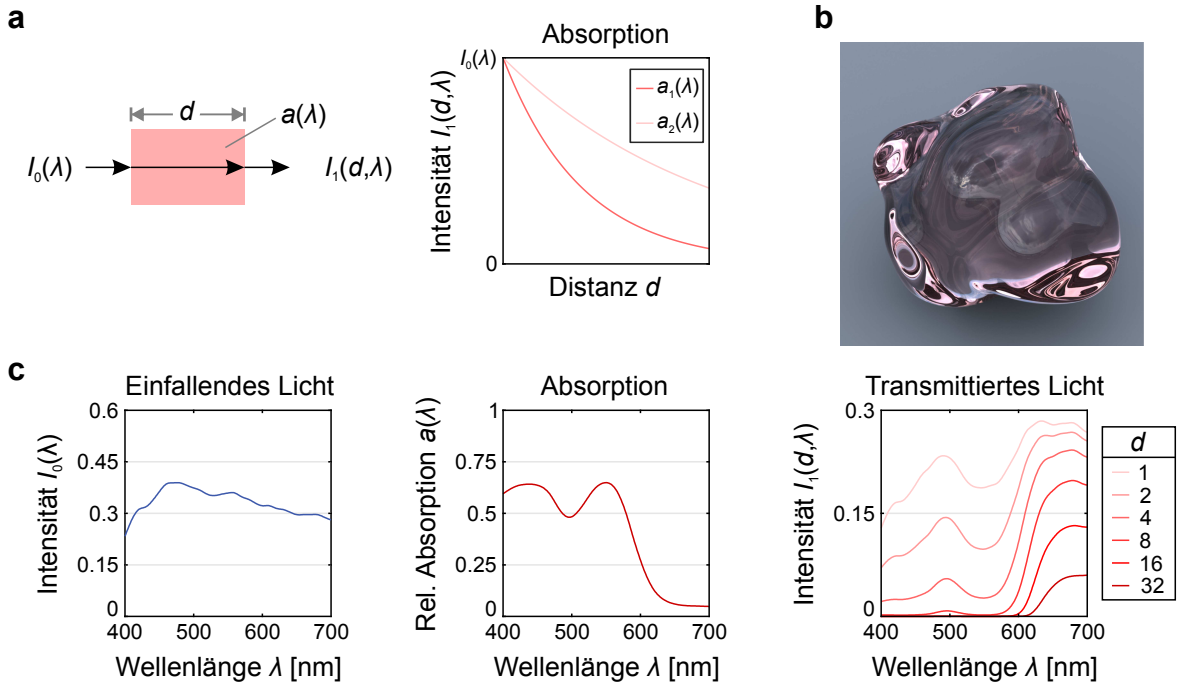

Abbildung 1.8: Absorption in lichtdurchlässigen Materialien. **(a)** Durchquert ein Lichtstrahl der Wellenlänge λ ein lichtabsorbierendes Material mit der Absorption $a(\lambda)$, reduziert sich seine ursprüngliche Intensität $I_0(\lambda)$ exponentiell mit der im Material zurückgelegten Distanz d zu $I_1(d, \lambda) = I_0(\lambda) \times \exp(-a(\lambda) \times d)$ (Bouguer-Lambert-Beersches Gesetz). In Materialien mit einem hohen Absorptionskoeffizienten ($a_1(\lambda)$, dunkelrote Kurve) nimmt die Intensität des Lichts schneller mit der zurückgelegten Distanz ab als in Materialien mit einem geringen Absorptionskoeffizienten ($a_2(\lambda)$, hellrote Kurve). **(b)** Sofern $a(\lambda)$ nicht für jede Wellenlänge λ denselben Wert hat, ändert sich neben der Intensität auch die relative spektrale Verteilung des transmittierten Lichts. Im hier gezeigten Beispiel wird kurzwelliges Licht stärker absorbiert als langwelliges. Als Folge wird die Intensität des langwelligen Lichts weniger stark abgeschwächt und das Objekt erscheint rötlich. **(c)** Veränderung der spektralen Energieverteilung eines Lichts durch Absorption am Beispiel von Tageslicht (linkes Diagramm; Taylor und Kerr, 1941, „sun plus sky“), das durch einen Farbfilter transmittiert wird, der kurzwelliges Licht stärker absorbiert als langwelliges (mittleres Diagramm für $d = 1$; Kodak CC50R, *Handbook of Kodak Photographic Filters*, 1990). Die Intensität des kurzwelligen Lichts nimmt mit zunehmender im Filter zurückgelegten Distanz d überproportional stark ab (rechtes Diagramm).

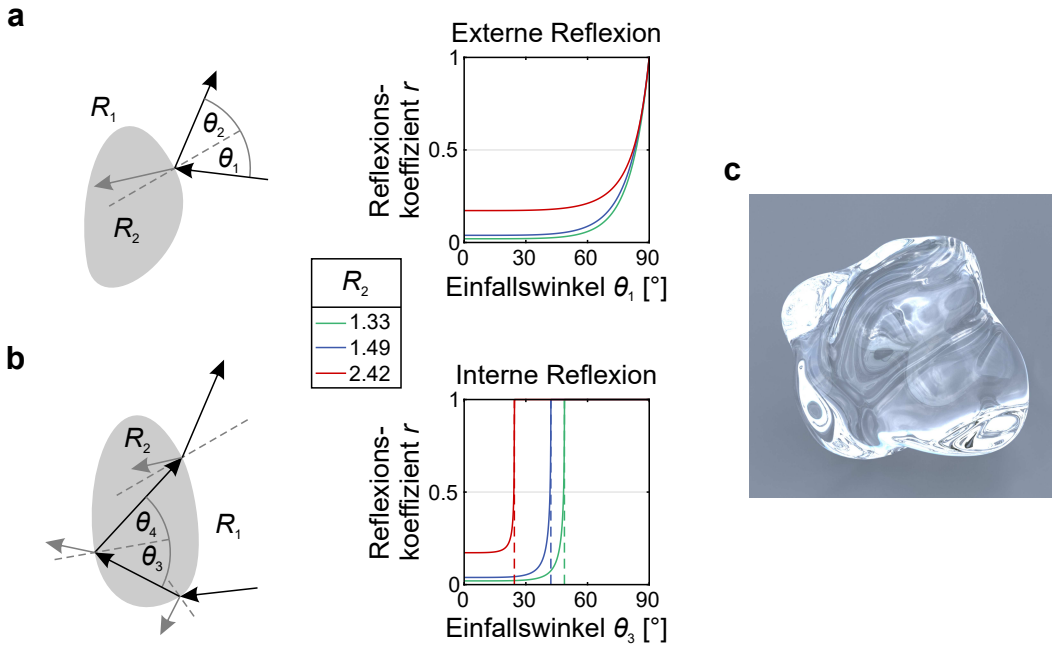

Abbildung 1.9: Reflexionen an lichtdurchlässigen Oberflächen. **(a)** Trifft Licht von außen auf ein lichtdurchlässiges Material, wird es zum Teil an seiner Oberfläche reflektiert. Sofern das Material eine glatte Mikrostruktur besitzt, ist der Reflexionswinkel θ_2 identisch zum Einfallswinkel θ_1. Wie groß der Anteil des reflektierten Lichts ist, wird durch die Fresnelschen Gleichungen beschrieben und hängt sowohl vom Brechungsindex des Materials (R_2) und dem seiner Umgebung (R_1) ab, als auch vom Einfallswinkel des Lichts (θ_1). Je größer R_2 und/oder θ_1, desto größer der Anteil des reflektierten Lichts. Dargestellt ist der Anteil des reflektierten Lichts (Reflexionskoeffizient r) für verschiedene von Luft ($R_1 \approx 1$) umgebene Materialien (Wasser: $R_1 = 1.33$; Glas: $R_1 = 1.49$; Diamant: $R_1 = 2.42$). **(b)** Trifft Licht von innen auf die Oberfläche eines lichtdurchlässigen Materials, wird es ebenfalls zum Teil reflektiert. Auch hier ist der Reflexionswinkel θ_4 identisch zum Einfallswinkel θ_3, sofern die Mikrostruktur der Oberfläche glatt ist. Dargestellt ist der Reflexionskoeffizient (r) für verschiedene von Luft umgebene Materialien. Wie in Abbildung 1.7c bereits erläutert, kann es an der Grenzfläche zum optisch dünneren Medium zur Totalreflexion kommen, sofern der Einfallswinkel einen bestimmten kritischen Wert überschreitet (senkrechte, gestrichelte Linien). **(c)** Die durch interne und externe spiegelnde Reflexionen verursachten Spiegelbilder der Umgebung sind im Bild additiv überlagert.

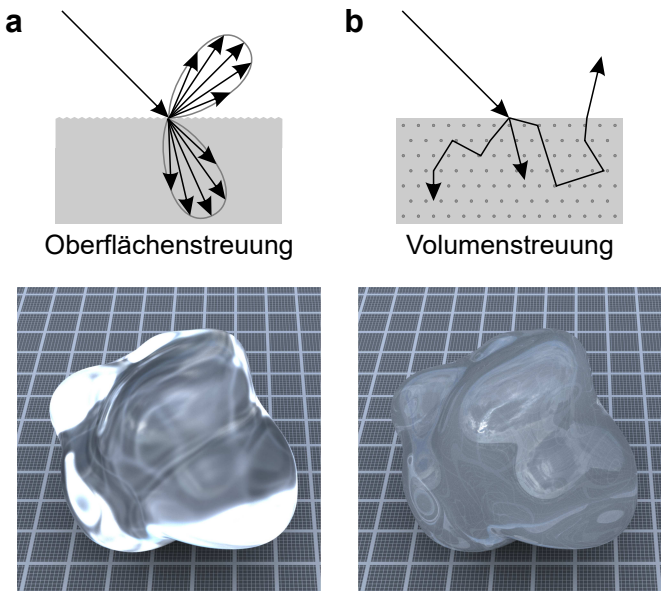

Abbildung 1.10: Oberflächen- und Volumenstreuung bei lichtdurchlässigen Materialien. **(a)** Die Mikrostruktur der Oberfläche lichtdurchlässiger Materialien bestimmt nicht nur, wie stark das an ihnen reflektierte Licht gestreut wird, sondern auch, wie stark das durch sie transmittierte und gebrochene Licht gestreut wird. Je rauer die Mikrostruktur der Oberfläche, desto ungerichteter erfolgen sowohl Reflexion als auch Transmission. **(b)** Zur Volumenstreuung kommt es, wenn Teile des durch eine lichtdurchlässige Oberfläche transmittierten Lichts im Inneren des Materials auf Partikel stoßen und von diesen reflektiert werden. Dabei können sich neben der Ausbreitungsrichtung des Lichts auch seine spektralen Eigenschaften verändern, sofern die Partikel Licht bestimmter Wellenlängen absorbieren.

lässigen Oberflächen darüber hinaus auch, wie stark das transmittierte Licht gestreut wird (siehe Abbildung 1.10a). Zusätzlich zu dieser nach innen gerichteten Oberflächenstreuung kann das transmittierte Licht auch innerhalb des Materials gestreut werden. Zu dieser sogenannten Volumenstreuung kommt es, wenn das Licht mit Partikeln im Inneren des Materials interagiert (siehe Abbildung 1.10b).

1.7.1 Unterscheidung zwischen Transparenz und Transluzenz

In Abhängigkeit von der Stärke der Streuungen an ihren Oberflächen und in ihrem Inneren lassen sich lichtdurchlässige Materialien grob in zwei Kategorien einteilen. Die erste Kategorie umfasst Materialien, die die räumliche Konfiguration des transmittierten Lichts durch Oberflächen- und/oder Volumenstreuungen so stark verändern, dass das ursprünglich vom Licht transportierte Bild nicht mehr als solches erkennbar ist. Wenn überhaupt, durchdringt nur ein kleiner Teil des Lichts diese Materialien vollständig und ohne Streuung. Solche Materialien erscheinen durchscheinend (transluzent). Zu dieser Kategorie von Materialien gehören zum Beispiel Wachs, Milch, Seife, Porzellan oder auch die menschliche Haut. Auch Luft kann transluzent sein, wenn sie durch kleine Staub- und/oder Flüssigkeitspartikel getrübt ist. Die

zweite Kategorie umfasst Materialien, deren Oberflächen- und/oder Volumenstreuung so gering ausgeprägt sind, dass das vom Licht transportierte Bild – abgesehen von der Modulation durch Lichtbrechung und Absorption – im Prinzip erhalten bleibt. Solche Materialien erscheinen durchsichtig (transparent). Zu dieser Kategorie von Materialien können zum Beispiel Wasser, Glas und Luft zählen, sofern sie hinreichend klar sind.

1.8 Informationen im retinalen Bild transparenter Materialien

Die vorhergehenden Betrachtungen zeigen, dass Licht mit lichtdurchlässigen Materialien zum Teil deutlich anders interagiert als mit opaken. Zwangsläufig hat dies nicht nur einen Einfluss auf den Informationsgehalt des von solchen Materialien zum Auge gelangenden Lichts, sondern auch darauf, mit welchen Regularitäten bestimmte Aspekte der Außenwelt im retinalen Bild verbunden sind. So enthält das von lichtdurchlässigen Materialien zum Auge gelangende Licht zum Beispiel Informationen über Aspekte, die im opaken Fall keinen direkten Einfluss auf den Lichttransport ausüben, wie zum Beispiel die Dicke eines Gegenstands, oder dort nur eine untergeordnete Rolle spielen, wie zum Beispiel die Brechungseigenschaften des Materials. Aus wahrnehmungspsychologischer Sicht erscheint insbesondere die Untermenge transparenter Materialien interessant, da in diesem Fall das vom Licht transportierte Bild im Prinzip erhalten bleibt und sich der Materialeinfluss im Wesentlichen auf die Modulation des bestehenden Bilds beschränkt. Die folgenden Betrachtungen zum Informationsgehalt des zum Auge gelangenden Lichts beziehen sich daher hauptsächlich auf solche transparenten Materialien.

Wie bereits erläutert, hängen im opaken Fall die Eigenschaften des Lichts oft wesentlich von den Eigenschaften der Oberfläche ab, mit der es zuletzt interagiert hat: Dabei führen ungerichtete Reflexionen nicht nur zu einer vollständigen Änderung der räumlichen Konfiguration des auftreffenden Lichts, sondern auch zu einer Mischung der bisherigen Einflüsse, denen das aus unterschiedlichen Richtungen stammende Licht ausgesetzt war. In der Folge transportiert das reflektierte Licht ein Bild, dessen Eigenschaften hauptsächlich durch die Eigenschaften der opaken Oberfläche selbst bestimmt werden. Diese Art des Informationseintrags durch diffuse Reflexion findet sich bei transparenten Materialien so nicht. In der Regel ist ihre Oberflächenstreuung nur gering ausgeprägt (sie wären sonst nicht transparent), sodass das an ihrer Oberfläche reflektierte Licht hauptsächlich aus der jeweiligen Spiegelungsrichtung stammt. Da die Geometrie dieser Spiegelungen grundsätzlich identisch ist zu der von opaken Materialien, erzeugen opake und lichtdurchlässige Oberflächen derselben Form und Mikrostruktur geometrisch identische Spiegelbilder. Auch die bei transparenten Oberflächen

zu beobachtende Abhängigkeit der Reflexionsstärke vom Einfallswinkel des Lichts findet sich im opaken Fall wieder, etwa bei Kunststoffen.

Anders als im opaken Fall, wo Reflexionen an der Oberfläche mit Absorption einhergehen können, verändern transparente Materialien die spektralen Eigenschaften des reflektierten Lichts nur aufgrund von Dispersion (vgl. Abbildung 1.7a). Das von transparenten Oberflächen an der ersten Grenzfläche direkt reflektierte Licht transportiert daher zwar ein Spiegelbild der Umgebung, enthält jedoch keine Informationen über die Absorptionseigenschaften des transparenten Materials.

Weitere entscheidende Unterschiede im Informationsgehalt basieren vor allem auf der Lichtdurchlässigkeit der Oberfläche selbst. Aufgrund dieser kommt von einem transparenten Objekt nicht nur Licht zum Auge, das von außen an seiner Oberfläche reflektiert wird, sondern auch Licht, das von hinten auf das Objekt trifft und durch dieses transmittiert wird. Wie die Reflexion erfolgt auch die Transmission aufgrund der geringen Oberflächen- und Volumenstreuung transparenter Objekte tendenziell gerichtet. Ähnlich wie im Fall der Spiegelungen bleibt die räumliche Konfiguration des transmittierten Lichts daher im Prinzip erhalten und damit auch das von ihm transportierte Bild. Im Unterschied zum außen reflektierten Licht, das den *vor* dem Objekt liegenden Teil der Szene abbildet (im Folgenden als „Umgebung“ bezeichnet), bildet das transmittierte Licht den *hinter* dem Objekt liegenden Teil der Szene ab (im Folgenden als „Hintergrund“ bezeichnet).

Wie das Spiegelbild, so kann auch das Bild des Hintergrunds mehr oder weniger stark optisch verzerrt werden. Das Ausmaß und die Art der Verzerrung hängen neben den Brechungseigenschaften des transparenten Materials auch von der Form seiner Oberfläche sowie den Eigenschaften des Hintergrunds ab und können daher im Prinzip Informationen über diese Aspekte liefern. Aufgrund der Lichtbrechung ist es im transparenten Fall allerdings nicht ohne Weiteres möglich, von einer Stelle des retinalen Bilds auf Eigenschaften der Außenwelt in einer bestimmten Richtung zu schließen. Das entsprechende Licht kann von ganz anderen Stellen der Szene stammen als von denen, aus deren Richtung es zum Auge gelangt.

Sofern das transmittierte Licht teilweise vom transparenten Material absorbiert wird, können sich zusätzlich zu den Verzerrungen seine Intensität und seine spektrale Verteilung ändern. Da diese Veränderungen auch von der Länge der Strecke abhängen, die das Licht innerhalb des Materials zurücklegt, liefern sie im Prinzip nicht nur Informationen über die Absorptionseigenschaften des Materials, sondern auch über seine Dicke.

Ein weiterer wesentlicher Unterschied zum opaken Fall ergibt sich aus der Tatsache, dass das von einem transparenten Material zum Auge gelangende Licht mehrmals auf verschiedene Weise mit dem Material interagiert haben kann. Dabei kann es von außen oder innen auf unterschiedliche Stellen seiner Oberfläche getroffen und dort reflektiert oder transmittiert (und gebrochen) worden sein und außerdem bei jeder Passage durch das Material teilweise absorbiert worden sein. Wie komplex sich das zum Auge gelangende Licht zusammensetzen

kann, wird zum Beispiel deutlich, wenn die Oberfläche eines transparenten Materials eine geschlossene Form besitzt, wie etwa bei einem vollständig von Luft umgebenen transparenten Objekt. Das vom Hintergrund durch ein solches Objekt transmittierte Licht wird in der Regel zwei Mal gebrochen. Einmal beim Eintritt in das Material, und ein weiteres Mal beim Austritt. Das durch den Eintritt in das Objekt optisch verzerrte Bild des Hintergrunds wird daher beim Austritt weiter verzerrt. Auf Bildebene betrachtet, sind diese beiden Verzerrungen nicht additiv, sondern hängen auf komplexe Weise voneinander ab, da jede Verzerrung die dreidimensionale Konfiguration des transmittierten Lichts verändert. Zusätzlich zu der komplexeren optischen Verzerrung des Hintergrunds kann bei solchen Objekten ein weiteres Spiegelbild der Umgebung an ihrer inneren Rückseite entstehen. Im Unterschied zu dem auf der Vorderseite gespiegelten Licht, interagiert das innen gespiegelte Lichtmuster mehrmals mit dem transparenten Material. Es wird in der Regel sowohl vor als auch nach der Reflexion vom Material teilweise absorbiert und an seiner Oberfläche gebrochen. Dieses innen auftretende Spiegelbild der Umgebung kann daher wesentlich stärker optisch verzerrt sein als das auf der vorderen Außenseite auftretende. Wie bei dem vom Hintergrund kommenden Licht hängt der Einfluss, den die Absorption auf das innere Spiegelbild ausübt, wiederum von der Dicke des Objekts ab.

Die Komplexität des Licht- und Bildtransports kann weiter zunehmen, wenn das zum Auge gelangende Licht auf seinem Weg mit weiteren lichtdurchlässigen Oberflächen interagiert. Dies ist zum Beispiel bei transparenten Objekten mit einem Hohlraum in ihrem Inneren der Fall. Bei ihnen wird das vom Hintergrund kommende Licht in der Regel vier Mal gebrochen und ist zwei Mal der Absorption durch das Objektmaterial ausgesetzt, bevor es zum Auge gelangt. Zusätzlich transportiert das zum Auge gelangende Licht bei solchen Objekten insgesamt vier Spiegelbilder der Umgebung, die an den verschiedenen inneren und äußeren Oberflächen des Objektmaterials entstehen. Das von innen an der äußeren Rückwand reflektierte Licht wird dabei auf seinem Weg in der Regel sogar sechs Mal gebrochen (drei Mal vor und drei Mal nach der Reflexion) und ist vier Mal der Absorption durch das Objektmaterial ausgesetzt (zwei Mal vor und zwei Mal nach der Reflexion), bevor es zum Auge gelangt.

Da im Prinzip beliebig viele transparente Oberflächen hintereinander angeordnet sein können, kann das zum Auge gelangende Licht im Prinzip unbegrenzt oft durch Lichtbrechung, Reflexion und Absorption beeinflusst sein. Diese Einflüsse sind in ihrer Wirkung insofern gleichwertig, als die Eigenschaften des transportierten Lichts nicht von der jeweils letzten Interaktion dominiert werden, wie es tendenziell bei diffus reflektierenden opaken Oberflächen der Fall ist. Insbesondere kommt es während der Interaktionen nie zu der für den opaken Fall beschriebenen vollständigen Neuformierung des vom Licht transportierten Bilds. Stattdessen werden die vom Licht transportierten Bilder fortlaufend mehr oder weniger stark in ihrer räumlichen Konfiguration, ihrer Intensität und ihren spektralen Eigenschaften verändert. Im

retinalen Bild können daher unzählige, mehrfach auf unterschiedliche Weise modulierte Bilder des Hintergrunds und der Umgebung überlagert sein.

1.9 Wahrnehmung transparenter Materialien

Die bisherigen Ausführungen zeigen, dass sich der Lichttransport transparenter und opaker Materialien und damit auch der Informationsgehalt des von ihnen kommenden Lichts deutlich voneinander unterscheiden. Aus den genannten Unterschieden ergeben sich mehrere, für diese Arbeit zentrale Fragen: (a) Wie funktioniert die visuelle Wahrnehmung im Fall transparenter Materialien? (b) Zu welchen Leistungen ist das visuelle Wahrnehmungssystem im transparenten Fall fähig? (c) Auf welche Regularitäten im retinalen Bild bezieht sich das visuelle System dabei? (d) Wie unterscheidet sich die Wahrnehmung im transparenten Fall von der im opaken?

1.9.1 Wahrnehmung einfacher transparenter Filter

Bisherige Arbeiten zur perzeptuellen Transparenz bezogen sich hauptsächlich auf die Farb- und Helligkeitsrelationen im retinalen Bild, die zum Auftreten von Transparenzeindrücken führen (Anderson, 2015; Beck, 1978; Beck, Prazdny & Ivry, 1984; Faul, 2017; Faul & Ekroll, 2002, 2011, 2012; Faul & Falkenberg, 2015; Kasrai & Kingdom, 2001; Khang & Zaidi, 2002a, 2002b; Metelli, 1970; Ripamonti, Westland & Da Pos, 2004; Robilotto, Khang & Zaidi, 2002; Singh & Anderson, 2002). Dabei wurden verhältnismäßig einfache Stimuli verwendet. Zum Teil bestanden diese nur aus wenigen jeweils homogenen Flächen unterschiedlicher Intensität. Diese vereinfachten Stimuli entsprechen in etwa den Bildern, die flache dünne Filter erzeugen, wenn sie sich vor einem ebenfalls flachen Hintergrund befinden, sie homogen beleuchtet und frontoparallel betrachtet werden (siehe Abbildung 1.11a). Dabei entsprechen die untersuchten Farb- und Helligkeitsrelationen den Modulationen des vom Hintergrund kommenden Lichts durch Absorption innerhalb des transparenten Materials. Andere Einflüsse transparenter Materialien auf den Lichttransport, wie etwa Spiegelungen, Lichtbrechung oder Streuung, treten in solchen Situationen entweder nur in vereinfachter Form zutage oder spielen nur eine untergeordnete Rolle. Untersucht wurde nicht nur, unter welchen Bedingungen perzeptuelle Transparenz auftritt, sondern zum Beispiel auch, wie das visuelle Wahrnehmungssystem auf Basis des retinalen Bilds die Farbe solcher Filter schätzt und wie robust solche Schätzungen gegenüber Veränderungen von Umgebungsfaktoren, wie etwa der Beleuchtung, sind.

1.9.2 Wahrnehmung komplexer transparenter Objekte

Im Unterschied zur Rolle von Farb- und Helligkeitsrelationen, wurde die Rolle, die optische Verzerrungen des Hintergrunds und Oberflächenspiegelungen für die Wahrnehmung

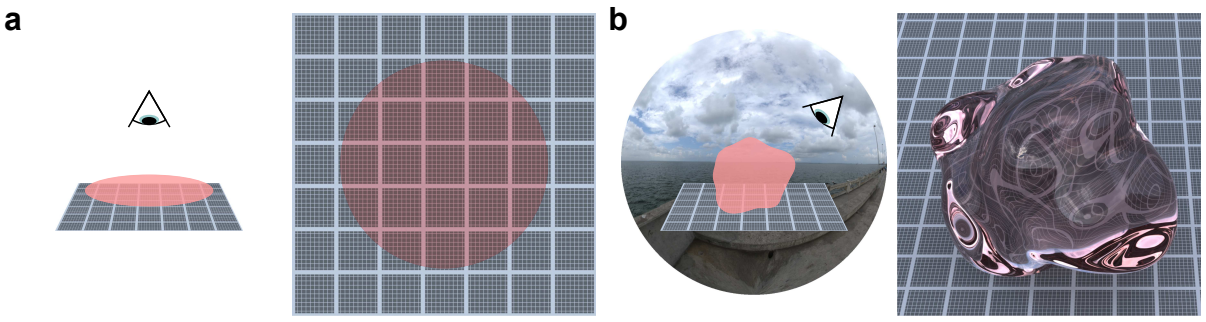

Abbildung 1.11: Beispiele für einfache und komplexe Situationen zur Untersuchung perzeptueller Transparenz. **(a)** Viele der Arbeiten zur Transparenzwahrnehmung verwenden verhältnismäßig einfache Stimuli. Diese entsprechen in etwa Situationen, in denen einfache flache Filter frontoparallel und unter homogener Beleuchtung betrachtet werden. Die Generierung entsprechender Stimuli geht teilweise mit Vereinfachungen des simulierten Lichttransports einher oder beschränkt sich sogar vollständig auf die Herstellung einfacher Farb- und Helligkeitsrelationen im Bild. **(b)** Bisher kaum untersucht wurde die Transparenzwahrnehmung in komplexeren Situationen mit realem (oder physikalisch plausibel simuliertem) Lichttransport, in denen dreidimensionale Objekte mit mehrfach gekrümmten Oberflächen einer realistischen Beleuchtung ausgesetzt sind und unter natürlichen Beobachtungsbedingungen betrachtet werden.

transparenter Materialien spielen, bisher kaum untersucht. Um sich mit diesen Einflüssen auseinandersetzen zu können, müssen in der Regel komplexere Situationen als die soeben beschriebenen herangezogen werden. Dazu zählen zum Beispiel dreidimensionale transparente Objekte mit einer mehrfach gekrümmten Oberfläche, die realistisch beleuchtet werden und unter natürlichen Beobachtungsbedingungen betrachtet werden (siehe Abbildung 1.11b). Beispiele, die uns im Alltag begegnen, sind Wasseroberflächen, Eiszapfen, Bernstein oder Gegenstände aus Glas oder Plastik, wie Flaschen oder Gläser. Diese komplexeren transparenten Objekte stellen den zentralen Untersuchungsgegenstand der folgenden Betrachtungen sowie der drei in dieser Arbeit vorgestellten Studien dar.

Unsere Alltagserfahrung mit komplexeren transparenten Materialien beziehungsweise Objekten ist in einem gewissen Sinn ambivalent. Auf der einen Seite haben wir in der Regel nicht den Eindruck, der Umgang mit transparenten Objekten sei problematischer als der mit opaken. Wir können transparente Materialien von opaken unterscheiden und können einen Eindruck ihrer Eigenschaften gewinnen. Neben ihrer Farbe nehmen wir in der Regel zum Beispiel auch die Form ihrer Oberfläche wahr. Dabei scheinen die in den einfacheren Situationen untersuchten Farb- und Helligkeitsveränderungen des transmittierten Lichts nicht einmal eine Voraussetzung für das Auftreten eines Transparenzeindrucks zu sein. Auch klare transparente Materialien ohne Absorption erscheinen uns transparent (wie etwa klares Wasser). Darüber hinaus können wir perzeptuell oft relativ gut die Eigenschaften eines transparenten Objekts von denen seines Hintergrunds trennen. Diese Wahrnehmungseindrücke scheinen zumindest

insoweit funktional zu sein, als sie uns ermöglichen, auch mit komplexeren transparenten Objekten in der Regel erfolgreich zu interagieren.

Auf der anderen Seite können im transparenten Fall Probleme auftreten, die uns aus dem Umgang mit opaken Materialien unbekannt sind. In solchen Fällen scheinen unsere Wahrnehmungseindrücke weniger funktional zu sein, als zunächst angenommen. Zum Beispiel kann die Beeinflussung des Lichttransports durch Lichtbrechung unter anderem zu inkonsistenten Formeindrücken führen. Wird ein (opaker) Gegenstand in eine transparente Flüssigkeit getaucht, so scheint es, als würde dieser an der Oberfläche der Flüssigkeit geknickt oder durchtrennt werden. Dieser Eindruck steht im Widerspruch zum Wissen, dass der Gegenstand seine Form durch diese Art der Manipulation eigentlich nicht verändern kann. Darüber hinaus kann es zum Beispiel schwierig sein, die Eigenschaften von Dingen wahrzunehmen, wenn diese nur durch transparente Materialien oder Objekte hindurch zu sehen sind und sie durch diese stark optisch verzerrt werden. Auch kann es schwerfallen, präzise nach Dingen zu greifen, wenn sich diese zum Beispiel *hinter* einem transparenten Objekt oder *in* einem transparenten Material befinden, wie zum Beispiel in Wasser. Greifen wir in derjenigen Richtung nach solchen Dingen, in der wir sie wahrnehmen, laufen wir aufgrund der Lichtbrechung Gefahr diese zu verfehlen. Entsprechende Situationen erfordern daher geeignete Strategien, um trotzdem angemessen mit der Umwelt interagieren zu können. So muss zum Beispiel beim Speerfischen absichtlich eine Stelle *vor* dem sichtbaren Ziel angepeilt werden, um den Einfluss der Lichtbrechung zu kompensieren.

Die genannten Beispiele zeigen, dass Situationen mit komplexeren transparenten Materialien und Objekten uns vor Probleme stellen können, die im opaken Fall, aber auch im Fall einfacher transparenter Filter, nicht auftreten. Stärker als in anderen Domänen der visuellen Wahrnehmung scheinen sich hier spezifische Grenzen der Leistungsfähigkeit des Wahrnehmungssystems zu offenbaren. Eine wissenschaftliche Auseinandersetzung mit der Wahrnehmung komplexer transparenter Objekte erscheint gerade deswegen vielversprechend, da spezifische Muster von Leistungen, Fehlleistungen oder fehlenden Leistungen den Raum möglicher Erklärungen für die Funktionsweise des Wahrnehmungssystems einschränken können und insofern besonders diagnostisch sein können. So ließe sich aus der im Alltag beobachtbaren Schwierigkeit, stark durch Lichtbrechung optisch verzerrte Objekte und ihre Eigenschaften wahrzunehmen zum Beispiel die Hypothese ableiten, dass das visuelle System grundsätzlich von einem geradlinigen Lichttransport ausgeht.

Die wissenschaftliche Auseinandersetzung mit der Transparenzwahrnehmung in komplexeren Situationen wurde lange Zeit durch den Umstand erschwert, dass die Simulation des Lichttransports in komplexen Szenen und die Synthese der entsprechenden retinalen Bilder mit immensem Aufwand verbunden war. Erst die jüngeren Entwicklungen im Bereich der computergestützten Bildsynthese erlauben es, komplexere Szenen mit dreidimensionalen transparenten Objekten mit vertretbarem Aufwand untersuchen zu können. Die heute übli-

chen Verfahren zur Bildsynthese approximieren den physikalischen Lichttransport zumindest so genau, dass sie physikalisch plausible Bilder generieren. Im Gegensatz zu früheren Algorithmen der Bildsynthese, berücksichtigen diese Verfahren zum Beispiel auch den Einfluss von Lichtbrechung sowie Volumenabsorption und -streuung. Mit Hilfe der modernen Verfahren ist es daher möglich, nahezu alle Einflussfaktoren komplexer Szenen auf das retinale Bild zu simulieren und zu manipulieren. Die computergestützte Bildsynthese stellt daher ein fundamentales Werkzeug zur Untersuchung der Transparenzwahrnehmung im Speziellen aber auch der visuellen Wahrnehmung im Allgemeinen dar.

Auf theoretischer Ebene hilft die Analyse der Bildsynthese, Regularitäten im Bild zu identifizieren, die Hinweis auf bestimmte Aspekte transparenter Objekte sein können. Dies geschieht zum Beispiel durch eine systematische Variation dieser Aspekte und Analyse der resultierenden Wirkung auf das Bild. Insbesondere für komplexe Phänomene, wie die Lichtbrechung, deren Einflüsse auf das Bild sich nicht unbedingt analytisch beschreiben lassen, stellt dies oft die einzige Möglichkeit zur Identifikation entsprechender Regularitäten dar.

Auf empirischer Ebene vereinfacht die Simulation der Bildsynthese, die Rolle bestimmter Bildregularitäten für die Wahrnehmung experimentell zu untersuchen. Durch systematische Manipulation des simulierten Lichttransports können zum Beispiel einzelne Regularitäten isoliert oder entfernt werden und so ihr Einfluss auf die Wahrnehmung untersucht werden. Der Lichttransport kann dabei nicht nur durch Veränderungen der physikalischen Eigenschaften der simulierten Szene beeinflusst werden, sondern auch durch direkte Eingriffe in die Simulation, indem zum Beispiel spezifische Lichtpfade während der Bildsynthese isoliert oder ignoriert werden.

1.9.3 Die Studie von Fleming, Jäkel und Maloney (2011b)

Eine der ersten Arbeiten, die sich mit der Transparenzwahrnehmung in komplexeren Situationen auseinandersetzt, stammt von Fleming, Jäkel und Maloney (2011b). Die Autoren wollten untersuchen, wie die Materialeigenschaften von dicken, irregulär geformten transparenten Objekten wahrgenommen werden. Im Gegensatz zu den bereits skizzierten Arbeiten zur Wahrnehmung flacher Filter, bezogen sich Fleming et al. (2011b) allerdings nicht auf die mit der Absorption zusammenhängenden Farb- und Helligkeitsrelation im Bild, sondern auf die Rolle der durch Lichtbrechung verursachten optischen Verzerrungen des Hintergrunds. Zu diesem Zweck betrachteten sie dicke Objekte mit mehrfach gekrümmten Oberflächen, deren Material keine Absorption besaß. Die Autoren stellten die Hypothese auf, dass der subjektive Materialeindruck, den solche klaren transparenten Objekte hervorrufen, wesentlich von ihrem Brechungsindex abhängt und dass der Brechungsindex vom visuellen System aus den optischen Hintergrundverzerrungen geschätzt werden kann (im Folgenden als „RI-Hypothese“ bezeichnet, mit „RI“ als Abkürzung für „refractive index“). Dazu würden in einem ersten

Schritt die optischen Verzerrungen aus dem retinalen Bild geschätzt und auf dieser Basis in einem zweiten Schritt der Brechungsindex des Objekts ermittelt werden. Laut Fleming et al. (2011b) könne sich das Wahrnehmungssystem zur Schätzung der Hintergrundverzerrungen auf das sogenannte Verzerrungsfeld („distortion field“) beziehen, welches die relativen Vergrößerungen und Verkleinerungen des Hintergrunds im Bild beschreibt. Die Autoren schlagen konkret vor, dass sich das Verzerrungsfeld durch einen Vergleich der relativen Größe von Texturelementen innerhalb des Objektbereichs und der Umgebung schätzen ließe. Das so geschätzte Verzerrungsfeld würde wichtige Aspekte der Divergenz des sogenannten Verschiebungsfelds („displacement field“) abbilden, welches die tatsächliche optische Verzerrung im Bild beschreibt. Die im geschätzten Verzerrungsfeld enthaltenen Informationen könnten den Autoren zufolge anschließend zu einer Schätzung des Brechungsindex integriert werden, zum Beispiel unter Bezug auf die mittlere Verzerrungsstärke.

Zum Test ihrer Hypothese führten Fleming et al. (2011b) ein Experiment durch, in dem sie Versuchspersonen die Materialien dicker, mehrfach gekrümmter transparenter Objekte abgleichen ließen. Dazu konnten die Versuchspersonen den Brechungsindex eines am Bildschirm gezeigten Objekts einstellen, bis das Objekt aus dem gleichen Material zu bestehen schien, wie ein gleichzeitig präsentiertes Vergleichsobjekt mit vorgegebenem Brechungsindex. Die beiden Objekte unterschieden sich entweder hinsichtlich ihrer Dicke oder ihres Abstand zum Hintergrund. Die von den Versuchspersonen eingestellten Brechungsindizes stimmten zwar ungefähr mit den vorgegebenen überein, allerdings gab es in Abhängigkeit von der Objektdicke und dem Hintergrundabstand größere systematische Abweichungen. Fleming et al. (2011b) schlossen daraus, dass die durch lichtbrechende Objekte verursachten Verzerrungen des Hintergrunds eine wichtige Informationsquelle darstellen, die das visuelle System nutzen würde, um die intrinsischen Materialeigenschaften transparenter Objekte zu schätzen.

Dass Fleming et al. (2011b) der Schätzung des physikalischen Brechungsindex eine zentrale Rolle bei der Wahrnehmung der Materialeigenschaften klarer transparenter Objekte beimessen ist insofern nachvollziehbar, als der Brechungsindex der einzige materialbezogene Parameter solcher Objekte ist, der die Hintergrundverzerrungen überhaupt beeinflusst. Es stellt sich allerdings die Frage, wie plausibel eine derartige Schätzung physikalischer Brechungseigenschaften durch das visuelle System tatsächlich ist und inwieweit die von Fleming et al. (2011b) gefundenen Ergebnisse tatsächlich als Indiz für eine derartige Schätzung interpretiert werden können. Um dies zu klären, setzen sich die ersten beiden in dieser Arbeit vorgestellten Studien intensiver mit der von Fleming et al. (2011b) vorgestellten RI-Hypothese, den von ihnen durchgeführten Experimenten sowie ihren Schlussfolgerungen auseinander.

1.10 Studie 1: Schätzung des Brechungsindex aus optischen Hintergrundverzerrungen

Die erste Studie (siehe Kapitel 2) analysiert die von Fleming et al. (2011b) aufgestellte Hypothese, das visuelle System würde die Brechungseigenschaften transparenter Objekte aus ihren Hintergrundverzerrungen schätzen, zunächst theoretisch aus computationaler und funktionaler Sicht und untersucht dann eine mögliche Alternativerklärung für die von den Autoren gefundenen Ergebnisse.

Aus computationaler Sicht ist die Schätzung des Brechungsindex aus den durch transparente Objekte verursachten optischen Hintergrundverzerrungen äußerst komplex, da die Verzerrungen nicht nur vom Brechungsindex, sondern auch von einer Reihe weiterer Faktoren, wie zum Beispiel vom Blickwinkel, der Objektform oder der Entfernung des Hintergrunds, abhängen. Dabei ist es nicht nur problematisch, dass die Einflüsse aller Faktoren im Bild des verzerrten Hintergrunds konfundiert sind. Einige dieser Einflussfaktoren können darüber hinaus auf eine Weise miteinander interagieren, dass sich ihre Einflüsse auf die räumliche Konfiguration des zum Auge gelangenden Lichts kompensieren. Zum Beispiel können zwei Objekte mit unterschiedlichen Dicken und Brechungseigenschaften quasi dieselben Verzerrungen des Hintergrunds hervorrufen.

Auch aus funktionaler Sicht ist die Plausibilität der von Fleming et al. (2011b) aufgestellten RI-Hypothese zumindest fragwürdig. Die meisten transparenten Materialien in unserer Umgebung sind entweder Wasser ($R \approx 1.33$) oder Glas ($R \approx 1.50$ bis 1.70). Andere Materialien sind deutlich seltener, wie zum Beispiel Diamant ($R \approx 2.42$). Sofern die Versuchspersonen von Fleming et al. (2011b) tatsächlich Materialabgleiche auf Basis von aus dem Bild geschätzten Brechungseigenschaften durchführten, wäre deren Genauigkeit im Verhältnis zur Bandbreite der vorkommenden Brechungsindizes verhältnismäßig gering. Eine derart ungenaue Schätzung und interne Repräsentation der Brechungseigenschaften transparenter Materialien könnte daher kaum zur funktionalen Kopplung an die Außenwelt beitragen. Unsere Alltagserfahrung zeigt, dass wir Materialien wie Wasser, Eis oder Glas anhand anderer Eigenschaften unterscheiden können, zum Beispiel anhand der Festigkeit oder der Temperatur.

Dass die RI-Hypothese aus computationaler und funktionaler Sicht unplausibel erscheint, schließt selbstverständlich nicht aus, dass das visuelles System trotzdem in der Lage sein könnte, die Brechungseigenschaften transparenter Objekte aus den von ihnen erzeugten Hintergrundverzerrungen zu schätzen. Sofern dies jedoch tatsächlich *nicht* der Fall ist, stellt sich die Frage, auf welcher Basis die Versuchspersonen von Fleming et al. (2011b) dann ihre Einstellungen vorgenommen haben. Eine plausible Alternativerklärung ist, dass die Versuchspersonen nicht geschätzte Brechungsindizes abgeglichen haben, sondern stattdessen versuchten, die Ähnlichkeit der gezeigten Stimuli auf Bildebene zu maximieren. In den von Fleming et al. (2011b) umgesetzten Situationen könnten sich solche Ähnlichkeitsabgleiche zum

Beispiel auf die mittlere Größe von Texturelementen des verzerrten Hintergrunds beziehen. Der verzerrte Hintergrund ist allerdings nicht das einzige Kriterium, das sich auf Bildebene abgleichen lässt. Wie bereits beschrieben, beeinflusst die Größe des Brechungsindex auch den Anteil des an Grenzflächen von Materialien reflektierten Lichts: Je höher der Brechungsindex, desto größer der Reflexionsgrad. Zusätzlich zur verzerrten Hintergrundtextur ließe sich daher zum Beispiel auch die Helligkeit der Oberflächenspiegelungen auf Bildebene abgleichen.

Um diese Alternativhypothese zu testen, wurde in der ersten Studie ein Abgleichsexperiment ähnlich zu dem von Fleming et al. (2011b) durchgeführt. Dabei wurden die Einstellungen der Versuchspersonen mit Vorhersagen für Ähnlichkeitsabgleiche der Hintergrundverzerrungen und der Spiegelungen verglichen. Diese Vorhersagen wurden zum Teil durch geeignete Simulationen, zum Teil empirisch gewonnen. Es stellte sich heraus, dass die Einstellungen der Versuchspersonen immer zwischen diesen beiden Vorhersagen lagen. War der Brechungsindex des vorgegebenen Objekts verhältnismäßig klein, entsprachen die Einstellungen eher denen, die bei einem Ähnlichkeitsabgleich der Hintergrundverzerrungen zu erwarten sind, bei größeren Brechungsindizes eher denen, die bei einem Ähnlichkeitsabgleich der Spiegelungen zu erwarten sind. Dieses Ergebnis spricht dafür, dass die Einstellungen der Versuchspersonen auf einem Kompromiss zwischen zwei bildbezogenen Ähnlichkeitskriterien basieren. Dieser Kompromiss scheint durch die Spiegelungen dominiert zu werden, sofern diese bei größeren Brechungsindizes hinreichend deutlich sichtbar werden. Dass sich die relativen Gewichte dieser beiden Ähnlichkeitskriterien für verschiedene Versuchspersonen zum Teil deutlich unterscheiden, kann als weiterer Hinweis darauf gewertet werden, dass das Einstellverhalten vornehmlich kognitiven Ursprungs ist und nicht auf einer Wahrnehmungsleistung beruht, die sich auf die Brechungseigenschaften der gezeigten Objekte bezieht.

1.11 Studie 2: Schätzung der optischen Hintergrundverzerrungen aus dem Bild

Die zweite Studie (siehe Kapitel 3) wählt einen anderen Zugang zur Auseinandersetzung mit der RI-Hypothese und beschäftigt sich mit der vorgeschlagenen Schätzung der optischen Hintergrundverzerrungen aus dem retinalen Bild. Laut Fleming et al. (2011b) würde das Verzerrungsfeld, welches die für die Schätzung des Brechungsindex relevanten Aspekte der Verzerrungen abbilden würde, durch einen Vergleich des durch das Objekt sichtbaren, verzerrten Teil des Hintergrunds mit dem unverdeckt sichtbaren, unverzerrten Teil geschätzt werden. Aus dieser Hypothese lassen sich zwei konkrete Vorhersagen ableiten, die in der zweiten Studie getestet wurden.

Die erste Vorhersage ist, dass Szenenvariablen, die das Verzerrungsfeld nicht verändern, wie die Dichte der Hintergrundtextur, die Einstellungen der Versuchspersonen in einer Mate-

rialabgleichsaufgabe nicht systematisch beeinflussen sollten. Um diese Hypothese zu testen, wurde ein Experiment ähnlich zu dem von Fleming et al. (2011b) durchgeführt, in dem Versuchspersonen das Material von transparenten Objekten abgleichen sollten, die sich, abgesehen von der Dichte ihrer Hintergrundtextur, in identischen Szenen befanden. Es stellte sich heraus, dass selbst relativ geringe Unterschiede in der Texturdichte des Hintergrunds einen systematischen und signifikanten Einfluss auf die Einstellungen der Versuchspersonen hatten. Da Änderungen der Texturdichte das Verzerrungsfeld nicht verändern, sprechen die gefundenen Abweichungen dafür, dass die Versuchspersonen ihre Einstellungen ohne den von Fleming et al. (2011b) vorgeschlagenen Bezug auf das Verzerrungsfeld durchführten.

Die zweite Vorhersage ist, dass die Unsicherheit etwaiger Brechungsindexschätzungen stark ansteigen sollte, wenn der unverdeckte Teil der Hintergrundtextur nicht für einen Vergleich zur Verfügung steht. Im Gegensatz dazu zeigten die Ergebnisse eines weiteren Experiments, dass sich die Einstellungen der Versuchspersonen nicht ändern, wenn die Hintergrundtextur außerhalb des Objekts durch eine homogene graue Textur ersetzt wird. Ein Vergleich von verzerrten und unverzerrten Hintergrundbereichen, wie er von Fleming et al. (2011b) vorgeschlagen wurde, scheint daher nicht stattzufinden.

In drei weiteren Experimenten wurde getestet, ob die gefundenen Ergebnismuster mit der in der ersten Studie vorgestellten alternativen Interpretation verträglich sind, dass Versuchspersonen schlicht die Ähnlichkeit der gezeigten Stimuli auf Bildebene maximieren. Dazu wurden die ersten beiden Experimente wiederholt, die Spiegelungen auf den Objektoberflächen jedoch entfernt. Dies isolierte nicht nur die in den Hintergrundverzerrungen enthaltenen Informationen, sondern erhöhte auch die Plausibilität von bildbezogenen Ähnlichkeitsabgleichen, da der Material- und Transparenzeindruck der so manipulierten Objektbilder reduziert war. Es stellte sich heraus, dass die eingestellten Brechungsindizes für solche Objekte zwar stärker von den vorgegebenen abwichen, die Systematik der Abweichungen jedoch dieselbe blieb wie in den ersten beiden Experimenten. Dieses Ergebnis ist mit der Interpretation vereinbar, dass die Versuchspersonen auch in den ersten beiden Experimenten bildbezogene Ähnlichkeitsabgleiche durchführten, ihre Einstellungen dort jedoch durch die Spiegelungen auf der Objektoberfläche in Richtung vermeintlich besserer Materialabgleiche verschoben wurden.

Darüber hinaus deuten Ergebnisse eines weiteren Experiments darauf hin, dass Versuchspersonen zwar im Prinzip dazu in der Lage sind, das Verhältnis zwischen verzerrten und unverzerrten Texturelementen des Hintergrunds zu berücksichtigen, sofern das experimentelle Design ihnen dies nahelegt, allerdings nur auf eine Weise, die nicht dazu geeignet ist, den Brechungsindex transparenter Materialien zu schätzen.

1.12 Fazit zu den ersten beiden Studien

Die in den ersten beiden Studien durchgeführte Analyse der Studie von Fleming et al. (2011b) ist nicht nur für das konkrete Thema der Transparenzwahrnehmung interessant, da sich an diesem Beispiel ganz allgemeine Probleme bei der wissenschaftlichen Untersuchung der Funktionsweise der menschlichen Wahrnehmung zeigen. Dazu gehört insbesondere die Gefahr, aufgrund von Methodenartefakten unzulässige Schlüsse zu ziehen. Ein Grund für derartige Fehlschlüsse können Probleme bei der Operationalisierung sein. Fleming et al. (2011b) haben die von Versuchspersonen eingestellten Brechungsindizes zur Messung ihrer Materialwahrnehmung verwendet. Dies ist allerdings nur dann eine geeignete Operationalisierung für die Materialwahrnehmung, wenn diese tatsächlich auf den Brechungseigenschaften der gezeigten Objekte beruht. Ist dies nicht der Fall, ist nicht ohne Weiteres klar, was mit den von den Versuchspersonen eingestellten Werten gemessen wird. Dieser Umstand wäre nur dann unproblematisch, wenn sich ein unerwartetes Vorgehen seitens der Versuchspersonen eindeutig als solches bemerkbar machen würde. Existiert jedoch ein alternatives Einstellverhalten, das zu Ergebnissen führen kann, die prinzipiell verwechselbar sind mit Ergebnissen, die aus einem hypothesenkonformen Einstellverhalten resultieren würden, ist die Gefahr von Fehlschlüssen besonders groß.

Darüber hinaus zeigt sich am Beispiel der Studie von Fleming et al. (2011b), wie auch der Einfluss etwaiger Störvariablen Fehlschlüsse zur Folge haben kann. Die in ihren Experimenten mit der Hintergrundverzerrung konfundierte Spiegelungsstärke war in dieser Hinsicht gleich in zweierlei Weise problematisch. Zum einen können die Spiegelungen als unkontrollierte Einflussgröße auf die mutmaßlichen Materialabgleiche aufgefasst werden, da sie, genau wie die im Fokus stehenden Hintergrundverzerrungen, ebenfalls von den Brechungseigenschaften transparenter Objekte abhängen. Zum anderen konnten die Spiegelungen in den von Fleming et al. (2011b) umgesetzten Situationen dazu führen, dass eingestellte und vorgegebene Brechungsindizes auch dann übereinstimmen, wenn die Versuchspersonen statt der mutmaßlichen Materialabgleiche lediglich Ähnlichkeitsabgleiche der Spiegelungshelligkeit durchführten. Ohne die nötige Sensibilisierung dafür, wie ein unerwartetes, alternatives Einstellverhalten seitens der Versuchsperson im Zusammenspiel mit Störvariablen wie den Spiegelungen zu vermeintlich hypothesenkonformen Messdaten führen kann, war die Gefahr besonders groß, die gefundenen Übereinstimmungen zwischen eingestellten und tatsächlichen Brechungsindizes fälschlicherweise als Beleg für die Gültigkeit der RI-Hypothese zu interpretieren.

An dieser Stelle zeigt sich außerdem, wie sehr computationale und funktionale Überlegungen helfen können, die Plausibilität einer Hypothese bereits im Vorfeld abzuschätzen. Gerade wenn derartige Überlegungen gegen die Plausibilität einer Hypothese sprechen, sind experimentell gewonnene Daten, die mit der aufgestellten Hypothese verträglich sind, mit ent-

sprechender Sorgfalt zu interpretieren und auf mögliche Methodenartefakte beziehungsweise Alternativerklärungen zu prüfen.

1.13 Studie 3: Visuelle Formwahrnehmung im Fall transparenter Objekte

Die in den ersten beiden Studien gewonnenen Hinweise darauf, dass die durch transparente Objekte verursachten optischen Verzerrungen des Hintergrunds nicht zur Schätzung ihrer Brechungsindizes genutzt werden, bedeutet nicht notwendigerweise, dass diese Verzerrungen für die Wahrnehmung transparenter Materialien keine Rolle spielen. Bereits unsere Alltagserfahrung gibt Hinweise darauf, dass die Erscheinung transparenter Materialien ganz wesentlich durch das Vorhandensein solcher Verzerrungen bestimmt wird. Auf Basis dieser Beobachtung wurde bereits in der ersten Studie angedeutet, dass optische Hintergrundverzerrungen etwa einen unspezifischen Hinweis auf das Vorhandensein eines transparenten Materials liefern könnten (siehe dazu auch Kawabe & Kogovšek, 2017; Kawabe, Maruya & Nishida, 2015; Kim & Marlow, 2016).

Die dritte in dieser Arbeit vorgestellte Studie (siehe Kapitel 4) setzt sich mit einer weiteren möglichen Rolle von optischen Hintergrundverzerrungen auseinander. Wie bereits skizziert, werden die Verzerrungen nicht nur durch den Brechungsindex eines transparenten Materials bestimmt, sondern auch durch andere Faktoren, wie etwa die Form seiner Oberfläche. Es ist daher nicht unplausibel anzunehmen, dass Hintergrundverzerrungen eine Bedeutung für die Formwahrnehmung haben könnten. Dies erscheint zumindest aus funktionaler Sicht plausibler als eine Schätzung der Brechungseigenschaften, da die Fähigkeit, die Form von Dingen wahrzunehmen, als wesentliche Voraussetzung dafür angesehen werden kann, erfolgreich mit der Umwelt zu interagieren.

Wie bereits erwähnt, ist die visuelle Formwahrnehmung Gegenstand einer Vielzahl von wissenschaftlichen Arbeiten. Fast alle dieser Arbeiten beziehen sich allerdings auf Objekte aus opaken Materialien. Eine auf den ersten Blick naheliegende Idee wäre, die bestehenden Befunde aus dem opaken Fall auf den Fall transparenter Oberflächen zu übertragen. Da sich der Lichttransport im transparenten Fall jedoch deutlich von dem im opaken Fall unterscheidet, ist dies nicht ohne Weiteres möglich. Bisher wurde die visuelle Formwahrnehmung lichtdurchlässiger Objekte nur wenig und oft in Bezug auf eng umgrenzte Fragestellungen untersucht (zum Beispiel Chen & Allison, 2013; Chowdhury, Marlow & Kim, 2017; Interrante, Fuchs & Pizer, 1995; Interrante, Fuchs & Pizer, 1997; Kersten, Stewart, Troje & Ellis, 2006). Es ist daher größtenteils ungeklärt, ob – und wenn ja, wie – sich die Formwahrnehmung transparenter Objekte von der opaker unterscheidet. Um diese Frage zu klären, setzt sich die dritte Studie sowohl theoretisch als auch empirisch mit der visuellen Formwahrnehmung im

Fall transparenter Objekte auseinander. Dabei liegt der Fokus nicht nur auf der Rolle der Hintergrundverzerrungen, sondern auch auf der Bedeutung weiterer mit der Form in Verbindung stehender Bildregularitäten. Spezifisch für transparente Objekte sind neben den Verzerrungen des Hintergrunds auch Farb- und Helligkeitsveränderungen durch Absorption und verzerrte Spiegelbilder der Umgebung, die durch spiegelnde Reflexionen an jeder Grenzfläche entstehen. In mehreren computationalen Analysen und Simulationen wird gezeigt, dass die Zusammenhänge zwischen diesen Bildregularitäten und der Form oftmals deutlich komplexer sind als bei den für opake Objekte vorgeschlagenen Formhinweisen. Darüber hinaus variiert die Komplexität des Zusammenhangs für verschiedene Situationen oft stark. Ist ein transparentes Objekt zum Beispiel in seinem Inneren hohl, verändert sich die Anzahl an Reflexionen und Lichtbrechungen, denen das zum Auge gelangende Licht ausgesetzt ist, und damit auch die Komplexität des Zusammenhangs zwischen den Bildregularitäten und der Form. Außerdem tritt das allgemeine Problem, dass verschiedene Bildregularitäten miteinander interagieren und daher nicht ohne Weiteres isoliert betrachtet werden können, im transparenten Fall in besonders deutlicher Form zutage. Die geschilderten Umstände legen den Schluss nahe, dass die Formwahrnehmung im Fall transparenter Objekte anders funktioniert als im Fall opaker Objekte.

Aus den vorgestellten theoretischen und computationalen Betrachtungen ergeben sich im Wesentlichen zwei Fragen: (a) Wie gut können wir die Form lichtdurchlässiger Objekte erkennen und wie gut gelingt dies verglichen mit opaken Objekten? (b) Welche der theoretisch analysierten Bildregularitäten spielen eine Rolle bei der Formwahrnehmung?

Um diese Fragen zu untersuchen, wurde ein Experiment durchgeführt, in dem Versuchspersonen computergenerierte Bilder von transparenten und opaken Objekten gezeigt wurden. Zur Messung ihrer Fähigkeit, die Form der Objekte zu erkennen, sollten die Versuchspersonen die lokale Oberflächenorientierung der Objekte mit Hilfe kleiner Messsonden angeben, die auf ihre Oberfläche projiziert wurden („gauge figure task“, Koenderink und van Doorn, 1992). Im Gegensatz zu einfachen Identifikationsaufgaben, bei denen Versuchspersonen ein bestimmtes Objekt unter mehreren Vergleichsobjekten identifizieren müssen, hat diese Normaleneinstellmethode den Vorteil, dass sie in der Regel nicht allein auf Basis der durch die Kontur bereitgestellten Forminformationen gelöst werden kann. Darüber hinaus erlaubt diese Methode nicht nur die Bildung lokaler Fehlermaße und damit die Analyse des Beitrags bestimmter lokaler Bildinformationen zur Formwahrnehmung, sondern auch die Rekonstruktion der wahrgenommenen Oberflächenform als Ganzes und damit die Bildung globaler Maße der Formerkennung (Nefs, 2008; Wijntjes, 2012).

Um zu untersuchen, ob die theoretisch analysierten Bildregularitäten eine Rolle bei der Formwahrnehmung transparenter Objekte spielen, wurde ihre Verfügbarkeit im retinalen Bild manipuliert. Dazu wurde in verschiedenen Bedingungen jeweils eine der Bildregularitäten weggelassen. Da die computationalen Analysen und Simulationen unter anderem gezeigt

haben, dass die Korrelation zwischen verschiedenen Bildregularitäten und der Form bei massiven und hohlen Objekten stark variieren kann, wurden im Experiment zwei verschiedene Varianten von transparenten Objekten verwendet: Entweder bestanden die Objekte vollständig aus dem transparenten Material oder sie waren dünnwandig und mit Luft gefüllt.

Die Ergebnisse des Experiments zeigen, dass die Form transparenter Objekte zwar wahrgenommen werden kann, die Einstellungen der Versuchspersonen jedoch sowohl weniger präzise als auch weniger korrekt waren als im opaken Fall. Außerdem zeigte sich, dass die im Experiment manipulierten Bildregularitäten die Formwahrnehmung beeinflusst haben. Der jeweilige Einfluss war im transparenten Fall jedoch teilweise gegensätzlich, je nachdem, ob die Regularitäten von massiven oder hohlen Objekten stammten. Die Ergebnisse deuten darauf hin, dass das visuelle System formbezogene Bildinformationen im transparenten Fall anders verarbeitet als im opaken Fall. Welche Informationen im Bild dabei als Formhinweis verwendet werden, scheint von einem komplexen Zusammenspiel der Eigenschaften des transparenten Objekts und der umgebenden Szene abzuhängen. Darüber hinaus ließen spezifische Muster lokaler Abweichungen der wahrgenommenen Form darauf schließen, dass manche Versuchspersonen in bestimmten Situationen Regularitäten im Bild, die mit der Form transparenter Objekte zusammenhängen, als Regularitäten opaker Objekte interpretiert haben. Dabei wurden Hintergrundverzerrungen durch Lichtbrechung und Farb- und Helligkeitsveränderungen durch Absorption offenbar wie die Textur beziehungsweise die Schattierung opaker Objekte interpretiert.

1.14 Gesamtfazit und Ausblick

Wenngleich die in dieser Arbeit vorgestellten Studien bereits verschiedene Rückschlüsse auf die Material- und Formwahrnehmung im Fall dreidimensionaler transparenter Objekte erlauben, stellen sie doch nur erste Schritte in einer systematischen Auseinandersetzung mit der Wahrnehmung dieser Klasse von Objekten dar. Zusätzlich zu den ausführlicheren Diskussionen, Schlussfolgerungen und Ausblicken innerhalb der drei Studien sollen daher im Folgenden kurz Strategien für das weitere Vorgehen skizziert werden, die für ein tiefer gehendes Verständnis perzeptueller Transparenz vielversprechend erscheinen. Neben Fragestellungen, die sich unmittelbar an die in den drei Studien vorgestellten Experimente anschließen, umfasst dies auch weiter gehende, allgemeinere Fragestellungen zur visuellen Wahrnehmung transparenter Materialien, auf die am Ende dieses Abschnitts eingegangen wird.

Bezogen auf die Materialwahrnehmung transparenter Objekte sprachen die in den ersten beiden Studien vorgestellten computationalen Überlegungen und empirischen Befunde gegen die von Fleming et al. (2011b) vorgeschlagene Hypothese, das Wahrnehmungssystem würde den Brechungsindex transparenter Materialien schätzen. Es ist jedoch nicht ausgeschlossen, dass das Wahrnehmungssystem unter optimalen Bedingungen trotzdem zu einer solchen

Schätzung in der Lage sein könnte. Um dies zu testen, müsste zunächst geklärt werden, welche Situationen im Hinblick auf eine derartige Schätzung eigentlich als optimal aufgefasst werden können. Dies könnte einfachere Situationen als die von Fleming et al. (2011b) umgesetzten ebenso umfassen wie komplexere Situationen, in denen zum Beispiel Teile der Szene bewegt sind. Außerdem müsste eine Experimentalsituation gefunden werden, in der Versuchspersonen entweder nicht auf einfache Bildabgleiche ausweichen können oder derartige Alternativstrategien zumindest eindeutig in den Einstellungen der Versuchspersonen identifiziert werden können. Ähnlichkeitsabgleiche der gezeigten Objekte auf Bildebene ließen sich zum Beispiel dadurch erschweren, dass Einstell- und Vergleichsobjekte nicht gleichzeitig, sondern nacheinander dargeboten werden. Darüber hinaus ließe sich die Zahl der potentiell auf Bildebene abgleichbaren Kriterien zum Beispiel durch die Wahl unterschiedlicher Beleuchtungen und/oder Hintergründe im Einstell- und Vergleichsreiz verringern. Eine weitere Möglichkeit bestünde darin, den geforderten Materialabgleich nicht mit Bezug auf einen explizit dargebotenen Vergleichsreiz durchzuführen, sondern mit Bezug auf einen möglicherweise existierenden internen Standard. Versuchspersonen könnten zum Beispiel dazu aufgefordert werden, den Brechungsindex eines dargebotenen Objekts so einzustellen, dass das Objekt aus einem bestimmten, ihnen bekannten Material zu bestehen scheint. Alternativ könnte statt einer aktiven Abgleichsaufgabe eine passive Identifikationsaufgabe verwendet werden. Versuchspersonen könnten zum Beispiel dazu aufgefordert werden zu entscheiden, ob eines von mehreren dargebotenen Vergleichsobjekten dasselbe Material besitzt wie ein vorgegebenes Objekt. Eine weitere mögliche Aufgabe könnte darin bestehen zu entscheiden, aus welchem Material ein unter verschiedenen Umgebungsbedingungen dargebotenes Objekt besteht.

Wie bereits diskutiert, ist der Brechungsindex zwar aus physikalischer Sicht ein entscheidender Parameter transparenter Materialien, er muss für die subjektive Materialwahrnehmung jedoch nicht notwendigerweise dieselbe Bedeutung besitzen. Die Auseinandersetzung mit der Materialwahrnehmung transparenter Objekte beschränkt sich daher nicht notwendigerweise nur auf die Frage, welche Rolle ihre Brechungseigenschaften für ihre Wahrnehmung spielen. Eine Aufgabe zukünftiger Arbeiten ist es daher, herauszuarbeiten, welche Dimensionen der subjektive Materialeindruck komplexer transparenter Objekte umfasst und wie diese mit ihren physikalischen Eigenschaften zusammenhängen. Einen naheliegenden Ausgangspunkt dafür bildet das von Faul und Ekroll (2002, 2011) und Faul (2017) vorgestellte Filtermodell, das die von flachen transparenten Filtern hervorgerufenen subjektiven Wahrnehmungseindrücke mittels verschiedener materialbezogener Dimensionen beschreibt.

Bezogen auf die Formwahrnehmung transparenter Objekte zeigt die dritte Studie, dass die Präzision und Korrektheit, mit der die Form dreidimensionaler transparenter Objekte wahrgenommen werden kann, vom Vorhandensein verschiedener Regularitäten im retinalen Bild abhängen. Dabei hatten manche Regularitäten einen unterschiedlich starken oder sogar

gegensätzlichen Einfluss, je nachdem ob sie durch ein massives oder ein hohles transparentes Objekt erzeugt wurden. Es ist daher nicht ohne Weiteres möglich, allgemeingültige Aussagen zur Rolle bestimmter Bildregularitäten bei der Formwahrnehmung dreidimensionaler transparenter Objekte zu tätigen. Stattdessen deuten die Ergebnisse darauf hin, dass die untersuchten Regularitäten nur in bestimmten Ausprägungen einen positiven Beitrag zur Formwahrnehmung leisten, in anderen hingegen keinen oder sogar einen negativen. Eine interessante Aufgabe zukünftiger Arbeiten wäre es, genauer zu untersuchen, von welchen Kriterien die potentielle Nutzbarkeit und die tatsächliche Nutzung der untersuchten Regularitäten als Formhinweis abhängt. Wie die dritte Studie zeigt, lassen sich diese Kriterien sowohl auf der Ebene der Bildregularitäten selbst beschreiben (zum Beispiel die Stärke der optischen Verzerrung im Bild) als auch auf der Ebene der Szeneneigenschaften (zum Beispiel die Anzahl lichtbrechender Grenzflächen, mit denen das zum Auge gelangende Licht interagiert). Dabei wäre es insbesondere interessant herauszufinden, wie sich die potentielle und tatsächliche Nutzung der Bildregularitäten über den Wertebereich dieser Kriterien hinweg verhalten. So ließe sich beispielsweise klären, ob sich der Beitrag einer Bildregularität zur Formwahrnehmung abrupt oder graduell ändert und ob der entsprechende Verlauf monoton ist oder nicht. Daran anknüpfend stellt sich die Frage, welchen Einfluss die jeweilige Nutzbarkeit eines Formhinweises auf die Art und Weise hat, mit der das visuelle System die aus ihm gewonnenen Informationen mit den aus anderen Formhinweisen gewonnenen Informationen integriert.

Zusätzlich zu den unterschiedlichen individuellen Einflüssen der untersuchten Bildregularitäten zeigt die dritte Studie, dass die Form dreidimensionaler transparenter Objekte insgesamt weniger präzise und korrekt erkannt wurde, als die opaker Objekte. Bei der Interpretation dieses Ergebnisses gilt es allerdings zu berücksichtigen, dass hier nur eine spezifische Situation untersucht wurde. Diese sollte auf der einen Seite verhältnismäßig komplex und realistisch sein, gleichzeitig jedoch erlauben, potentielle Unterschiede in der Formwahrnehmung auf die manipulierten Bildregularitäten zurückzuführen. Wie bereits beschrieben, wurden dazu unter anderem physikalisch plausible Objekte und Szenen verwendet und diese stereoskopisch dargeboten, gleichzeitig jedoch zufällige, unbekannte Formen verwendet und zahlreiche Szeneneigenschaften, wie die Beleuchtung und die Beobachterposition, konstant gehalten. Wie bei jeder experimentellen Untersuchung stellt sich deshalb die Frage, inwieweit sich die hier gefundenen Ergebnisse auf weniger restriktive Situationen verallgemeinern lassen.

Einige Indizien sprechen dafür, dass in der dritten Studie eine aus Wahrnehmungssicht verhältnismäßig ungünstige Situation umgesetzt wurde und die Formwahrnehmung transparenter Objekte in anderen Situationen besser ausfallen könnte. Zum Beispiel haben die vorgestellten computationalen Betrachtungen gezeigt, dass sich die Form einer Oberfläche verhältnismäßig einfach aus ihren optischen Verzerrungen des Hintergrunds schätzen ließe, wenn das vom Hintergrund kommende Licht nur einmal gebrochen wird und nicht mehrmals, wie bei den im Experiment verwendeten Objekten. Es wäre daher interessant empirisch zu

prüfen, ob die Form in solchen weniger komplexen Situationen, wie zum Beispiel beim Blick auf eine gewellte Wasseroberfläche, besser erkannt werden kann und, falls ja, wie weit sich das Fehlerniveau dem des opaken Falls annähert. Solche einfacheren Situationen bieten darüber hinaus den Vorteil, dass sich der in der dritten Studie beschriebene Effekt der Vergrößerungsumkehr („magnification inversion"), der die Eindeutigkeit des Zusammenhangs zwischen optischen Vergrößerungen und der Oberflächenform auflöst, besser untersuchen ließe: Anders als im Fall komplexer dreidimensionaler Objekte mit geschlossener Form lässt sich das Auftreten der Vergrößerungsumkehr für einzelne lichtbrechende Grenzflächen, wie etwa der genannten Wasseroberfläche, gezielter experimentell manipulieren. Auch die in der dritten Studie diskutierten Interaktionen zwischen verschiedenen mit der Form zusammenhängenden Bildregularitäten ließe sich in einfacheren Situationen experimentell besser kontrollieren als in komplexen, da das Licht insgesamt weniger oft mit dem Material und seiner Oberfläche interagiert.

Nicht nur eine Verringerung der Komplexität, sondern auch eine Anreicherung der betrachteten Situation könnte zu einer Verbesserung der Formwahrnehmung führen. Der Grund dafür ist, dass unter realistischeren Beobachtungsbedingungen weitere mit der Form in Verbindung stehende Bildregularitäten wirksam werden können, die in der bisher betrachteten Situation entweder konstant gehalten wurden oder gar nicht vorhanden waren. Wie in der dritten Studie bereits angedeutet, gehören dazu unter anderem Bewegungen von Objekt und/oder Beobachter. Da das visuelle System die aus Bewegungen resultierenden Regularitäten offenbar dazu nutzt, um transparente Materialien als solche zu identifizieren (Kawabe & Kogovšek, 2017; Kawabe et al., 2015; Tamura, Higashi & Nakauchi, 2018), ist es nicht unplausibel anzunehmen, dass diese Regularitäten auch bei anderen Wahrnehmungsleistungen wie der Formwahrnehmung eine Rolle spielen. Die Auseinandersetzung mit dieser Frage ist insofern besonders anspruchsvoll, als die Identifikation und Beschreibung von mit der Form zusammenhängenden Bildregularitäten im dynamischen Fall deutlich komplexer ist als im statischen. Im bewegten Fall können sich Informationen nicht nur aus der räumlichen Konfiguration des retinalen Bilds ergeben, sondern zusätzlich auch aus seiner zeitlichen. Dass durch Bewegung verursachte Bildregularitäten im Prinzip dazu verwendet werden könnten, die Form transparenter Objekte zu schätzen, wurde zum Beispiel bereits von Ben-Ezra und Nayar (2003) näher untersucht.

Eine weitere Möglichkeit, die betrachtete Situation anzureichern, besteht darin, die Unvollkommenheiten natürlicher Oberflächen stärker zu berücksichtigen. Dazu zählen zum Beispiel kleine Unebenheiten, Abnutzungen, Verschmutzungen und Kratzer. In der computergestützten Bildsynthese werden solche Unvollkommenheiten vor allem deswegen umgesetzt, weil sie die wahrgenommene Realitätsnähe der Bilder erhöhen (siehe zum Beispiel Becket und Badler, 1990; Schwärzler und Wimmer, 2007). Darüber hinaus könnten sie auch für die Formwahrnehmung eine Rolle spielen: Befinden sich auf einer transparenten Oberfläche zum Beispiel Staubpartikel, erzeugen diese eine Art Textur, die sich im Bild ähnlich verhält wie die

Textur diffus reflektierender opaker Oberflächen. Es wäre daher interessant zu untersuchen, inwieweit solche unter realen Beobachtungsbedingungen auftretenden Unvollkommenheiten die in der dritten Studie gefundene schlechtere Formwahrnehmung im transparenten Fall kompensieren könnten.

Im Rahmen der dritten Studie durchgeführte informelle Voruntersuchungen deuten darauf hin, dass die Formwahrnehmung nicht nur von der jeweils betrachteten Situation, sondern auch von der Art ihrer Operationalisierung abhängt. Beispielsweise scheint die Leistung in der Aufgabe, die Gleichheit zweier Objekte anhand ihrer Form zu bewerten, tendenziell weniger stark von ihrem Material abzuhängen, als die Aufgabe ihre lokale Oberflächenausrichtung anzugeben. Es scheint daher formbezogene Aufgaben zu geben, deren Materialabhängigkeit geringer ausgeprägt ist und es bleibt zu klären, wodurch diese charakterisiert sind. Die Befunde aus den Voruntersuchungen können als ein erster Hinweis darauf gewertet werden, dass Operationalisierungen, die stärker an natürliche Interaktionen mit unserer Umwelt angelehnt sind, tendenziell weniger abhängig vom Material sind. Ein möglicher Grund für eine bessere Formwahrnehmung bei natürlicheren Aufgaben könnte darin bestehen, dass sich sowohl die Menge und Art der zur Bewältigung der Aufgabe benötigten Informationen als auch ihre Verfügbarkeit in einer gegebenen Situation von derjenigen bei weniger natürlichen Aufgaben unterscheidet. So scheint zum Beispiel die äußere Kontur eines Objekts auszureichen, um zu entscheiden, ob sich seine Form insgesamt von der eines anderen Objekts unterscheidet, während der Nutzen der Kontur für die Angabe lokaler Oberflächenausrichtungen mit zunehmendem Abstand vom Rand deutlich abnimmt.

Insbesondere wenn die Formwahrnehmung über Aufgaben operationalisiert wird, die mehrere Sinne umfassen, nimmt die Menge der zur Verfügung stehenden Informationen zu. Dies ist zum Beispiel der Fall, wenn ein Objekt nicht nur angesehen, sondern auch angefasst werden kann. Dem Wahrnehmungssystem stehen dann zusätzlich zu den visuellen auch haptische Informationen über die Form zur Verfügung. Wie diese haptischen Informationen für die Formwahrnehmung genutzt und mit visuellen Informationen integriert werden können, wurde bereits in einigen Arbeiten genauer untersucht (zum Beispiel Gaissert und Wallraven, 2012; Norman, Norman, Clayton, Lianekhammy und Zielke, 2004a; Norman et al., 2012; Phillips, Egan und Perry, 2009). Da in diesen Arbeiten jedoch nur opake Objekte verwendet wurden, bleibt es zu klären, inwieweit sich die entsprechenden Befunde auf Situationen übertragen lassen, in denen die visuellen Informationen von transparenten Objekten stammen. Dabei wäre es insbesondere interessant zu untersuchen, auf welche Weise die Integration erfolgt, wenn es bei der visuellen Formschätzung zu systematischen Fehleinschätzungen kommt, wie zum Beispiel bei den in der dritten Studie verwendeten massiven transparenten Objekten.

Allgemein stellt sich nicht nur die Frage, auf Basis welcher Mechanismen die Wahrnehmung von Material- und Formeigenschaften transparenter Objekte geschieht, sondern auch, wovon die Aktivierung dieser transparenzspezifischen Mechanismen überhaupt abhängt. Die

Bedingungen perzeptueller Transparenz wurden nicht nur für flache Filter, sondern auch für komplexe dreidimensionale Objekte bereits genauer untersucht, bei letzteren insbesondere mit Bezug auf die Rolle der Hintergrundverzerrungen (Kawabe & Kogovšek, 2017; Kawabe et al., 2015; Kim & Marlow, 2016; Tamura et al., 2018). Die in der dritten Studie durchgeführte Nachbefragung zeigt, wie das Entstehen perzeptueller Transparenz darüber hinaus auch vom Vorhandensein weiterer Bildregularitäten abhängt. Zum Beispiel reduzierte die Wegnahme von Oberflächenspiegelungen die Anzahl derjenigen Personen, die ein transparentes Objekt auch als solches wahrnahmen, insbesondere dann, wenn das Objekt massiv war und seine Hintergrundverzerrungen entsprechend stark ausgeprägt waren. Die physikalischen und perzeptuellen Materialkategorien scheinen daher nicht grundsätzlich übereinzustimmen. Insbesondere ist das Vorliegen physikalischer Transparenz (beziehungsweise die Simulation der durch sie erzeugten Bilder) keine hinreichende Bedingung für perzeptuelle Transparenz. Eine mögliche Schlussfolgerung ist, dass die perzeptuellen Grenzen zwischen verschiedenen Materialien im Vergleich zu den physikalischen Grenzen verschoben sind. Dabei gilt es jedoch zunächst zu bedenken, dass selbst auf physikalischer Ebene keine eindeutigen Grenzen zwischen opaken, transluzenten und transparenten Materialien definiert sind (vgl. Abschnitt 1.7.1). Stattdessen sind die Übergange zwischen den verschiedenen Materialkategorien fließend und die vermeintlich objektive Zuordnung eines Materials zu einer der Kategorien beruht üblicherweise selbst wiederum auf subjektiven Kriterien, wie der Erkennbarkeit von Gegenständen hinter dem Material. Von dieser Problematik abgesehen zeigt die Nachbefragung, dass die Übergange zwischen den perzeptuellen Materialkategorien teilweise so weit verschoben sein können, dass selbst der aus physikalischer Sicht eindeutige Fall eines vollständig transparenten Objekts von manchen Personen fälschlicherweise als opak eingeschätzt wurde. Ein weiteres Problem ergibt sich aus der Tatsache, dass der bewusst verfügbare Materialeindruck nicht notwendigerweise widerspiegeln muss, welche materialspezifischen Mechanismen während des Wahrnehmungsprozesses tatsächlich aktiviert werden. Auf die Aktivierung transparenzspezifischer Mechanismen kann daher nur indirekt geschlossen werden, was deren Untersuchung entsprechend erschwert. Dessen ungeachtet ließe sich durch ein genaueres Verständnis der Bedingungen perzeptueller Transparenz im Fall komplexer transparenter Objekte nicht nur weiter eingrenzen, von welchen Bildregularitäten die Aktivierung transparenzspezifischer Mechanismen abhängt, sondern auch, auf welche Weise transparente Materialien und ihre Eigenschaften vom Wahrnehmungssystem intern repräsentiert werden.

Kapitel 2

Are optical distortions used as a cue for material properties of thick transparent objects?

Dieses Kapitel wurde veröffentlicht als: Schlüter, N. & Faul, F. (2014). Are optical distortions used as a cue for material properties of thick transparent objects? *Journal of Vision*, *14*(14), 1–14. https://doi.org/10.1167/14.14.2

Zusammenfassung

Fleming, Jäkel und Maloney (2011b) ließen Versuchspersonen die wahrgenommenen Materialeigenschaften von dicken, klaren transparenten Objekten in fotorealistischen Szenen durch Einstellungen des Brechungsindex abgleichen. Sie fanden ungefähre Übereinstimmungen zwischen Standard- und Testobjekten, aber auch große systematische Abweichungen. Dennoch kamen sie zu dem Schluss, dass die geschätzte Brechung als Indikator für die Materialeigenschaften von transparenten Objekten verwendet wird und betonten die Rolle objektinduzierter Hintergrundverzerrungen in diesem Prozess. Dies erscheint jedoch nicht plausibel, da die notwendigen Informationen zur Schätzung des Brechungsindex aus Verzerrungen – zum Beispiel die genaue Form und Dicke des Objekts, sein Abstand zum Hintergrund und der unverzerrte Hintergrund – in der von ihnen umgesetzten Situation nicht verfügbar waren. Eine plausiblere Alternativerklärung ist, dass die Versuchspersonen nicht geschätzte Brechungsindizes abgeglichen haben, sondern einfache Ähnlichkeitsabgleiche von Bildattributen durchführten, die sich aus den Hintergrundverzerrungen oder den Spiegelungen ergeben. Wir testeten diese Hypothese in einem ähnlichen Abgleichsexperiment, in dem es möglich war, sowohl den Brechungsindex für einen Ähnlichkeitsabgleich der Hintergrundverzerrungen als auch den für einen Ähnlichkeitsabgleich der spiegelnden Reflexionen vorherzusagen. Unsere Versuchspersonen wählten immer einen Wert zwischen diesen beiden Vorhersagen. Die spiegelnde Reflexion ist dabei tendenziell der dominierende Faktor, sobald sie deutlich sichtbar wird.

Unsere Ergebnisse sind verträglich mit der Annahme, dass die Versuchspersonen versucht haben, einen Kompromiss zwischen zwei bildbasierten Ähnlichkeitskriterien zu finden. Sie scheinen nicht mit der Annahme vereinbar zu sein, dass die Abgleiche auf der Grundlage interner Schätzungen der Brechungsindizes erfolgen.

Abstract

Fleming, Jäkel, and Maloney (2011b) asked subjects to match perceived material properties of thick, clear transparent objects in photo-realistic scenes by adjusting the refractive index. They found approximate correspondence between standard and test objects but also large systematic deviations. Nevertheless, they concluded that estimated refraction is used as an indicator for material properties of light-transmitting objects and emphasized the role of object-induced background distortions in this process. This, however, seems not plausible, because the necessary information for inferring the refractive index from distortions – for example, the object's exact shape and thickness, its background distance, and the undistorted background – was not available in their situation. A more plausible alternative explanation is that the subjects did not match estimated refractive indices, but instead performed simple similarity matches based directly on image attributes related to background distortions or specular reflections. We tested this hypothesis in a similar matching experiment in which it was possible to predict the refractive index for a similarity match based on background distortions and for a similarity match based on specular reflections. Our subjects always chose a value between these two predictions. The specular reflection tends to be the dominant factor as soon as it becomes clearly noticeable. Our findings are compatible with the assumption that the subjects tried to find a compromise between two image-based similarity criteria. They do not seem to be consistent with the assumption that the matches are made on the basis of internal estimates of refractive indices.

2.1 Introduction

Our daily experience shows that we are able to perceive material properties of objects solely on the basis of visual information. We can effortlessly tell if objects are soft or hard, stable or fragile, rough or smooth. Objects that transmit light, like amber or icicles, are especially fascinating, because they impress us with their play of colors, distortions, and reflections. From a theoretical point of view, such transparent objects are especially interesting because they present the visual system with problems that do not occur with opaque objects: The retinal images of opaque objects depend mainly on their intrinsic material properties and the prevailing illumination, whereas those of transparent objects are also influenced by properties of the background that is visible through them. This raises the question, to what extent the

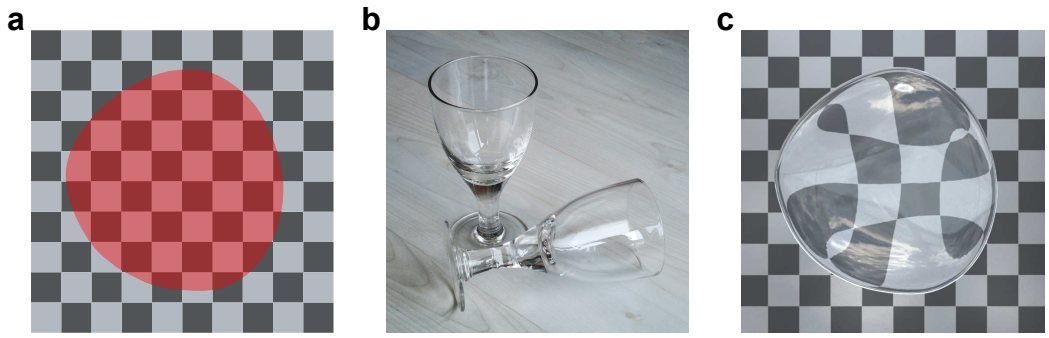

Figure 2.1: Different types of transparent objects. **(a)** Flat transparent objects, viewed frontally under a homogeneous illumination are well-suited for investigating the perception of transmission-related material properties. **(b)** More complex transparent objects reveal that the impression of transparency does not only depend on transmission-related properties, but also on other material characteristics, for example, the refractive properties. **(c)** Massive, relatively thick objects with curved surfaces without internal absorption or self-occlusion are well-suited for isolating refraction-related properties.

visual system uses information related to the background to infer material properties of transparent objects.

Historically, investigations of perceptual transparency arose from work on color perception. It is therefore not surprising that mainly proximal color and brightness regularities were examined as cues for transparent objects (Metelli, 1970; Singh & Anderson, 2002). In filter models of perceptual transparency, potential color-related regularities were derived from the spectral transmission properties of optical filters (Beck, 1978; Beck, Prazdny, & Ivry, 1984; Faul & Ekroll, 2002, 2011). To investigate transmission-related properties that lead to color changes of the background in isolation, highly simplified situations were used in which flat transparent filters were viewed frontally under a homogeneous illumination (see Figure 2.1a). If one considers more complex objects under more natural viewing conditions, then it is obvious that the transparency impression depends on additional material properties. This is, for instance, demonstrated by highly transmissive objects like a colorless wine glass that changes the spectral composition of the transmitted light only slightly but can nevertheless elicit strong impressions of transparency (see Figure 2.1b). Potential cues that may be used to detect transparency in these cases may relate to the fact that light-transmitting objects distort the transmitted background texture and specularly reflect the environment on their surface. Both characteristics depend directly on a material property, the refractive index, and are therefore a natural basis from which properties of transparent objects may be inferred. It is important to note, however, that the refractive index is but one of several factors that influence distortions and reflections.

Fleming, Jäkel, and Maloney (2011b) were the first to systematically investigate the material perception of such complex transparent objects. In particular, they tested the hypothesis that the optical distortion of the background is used to infer refractive material properties of thick,

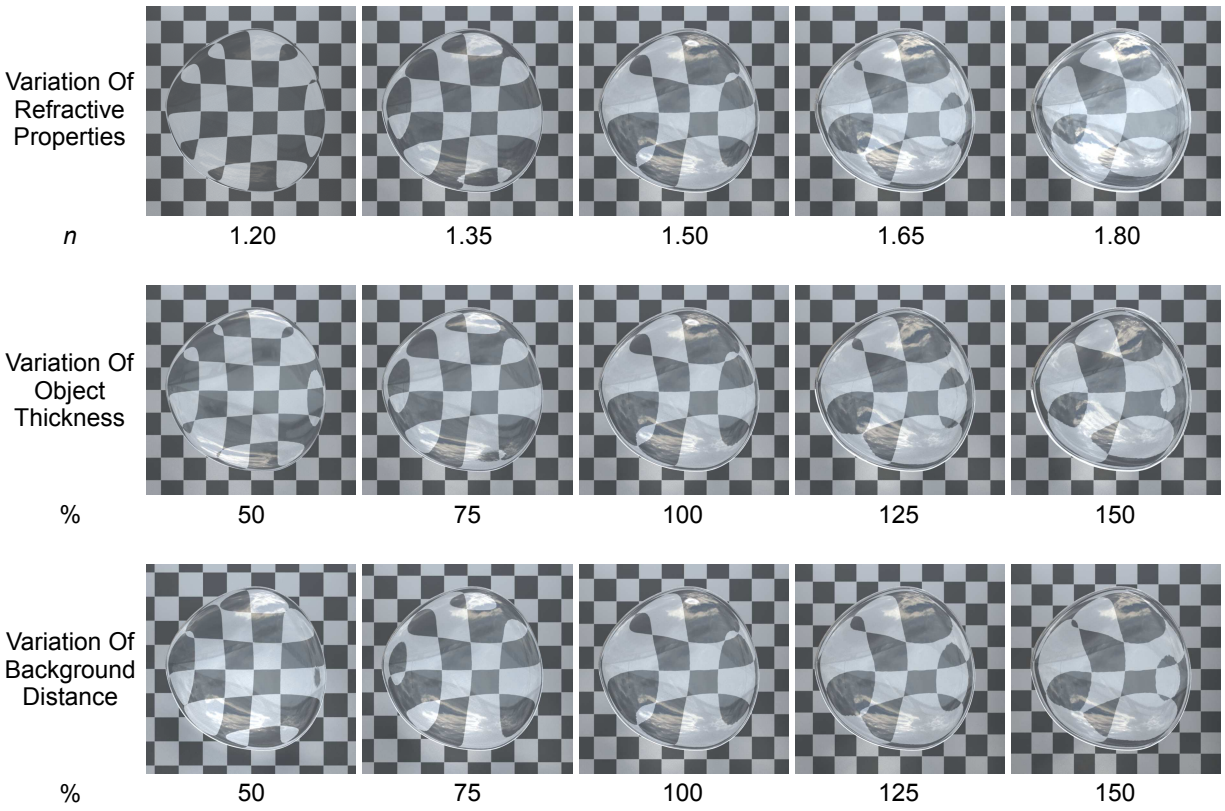

Figure 2.2: The optical distortion of the background is not only influenced by the refractive index of the object material (first row, object thickness and background distance are 100 %, i.e., 6 mm and 60 mm, respectively), but also by context factors like the object thickness (second row, refractive index = 1.5, background distance = 100 %) or the background distance (third row, refractive index = 1.5, object thickness = 100 %).

clear transparent objects. This hypothesis seems to be reasonable, because the refractive index is the only material-related property that influences the distortions of the background texture in the proximal image. To isolate refraction-related properties, they used simulated images of massive, relatively thick objects with curved surfaces that had an internal absorption of zero (see Figure 2.1c for an exemplary depiction of such objects). In each trial of their matching experiments, they presented two such objects in different static scenes and the subject's task was to adjust the refractive index of the test object until it appeared to be made of the same material as the standard object.

Such matches are nontrivial because the optical distortion of the background does not only depend on the refractive index of the transparent object but also on numerous context factors (e.g., object thickness, object shape, background distance, Figure 2.2). If the context of the two presented objects differs in at least one of these factors, then it is not possible to match the refractive properties on the basis of a simple proximal match. In their experiments, Fleming et al. (2011b) varied the object thickness and the background distance (cf. Figure 2.3).

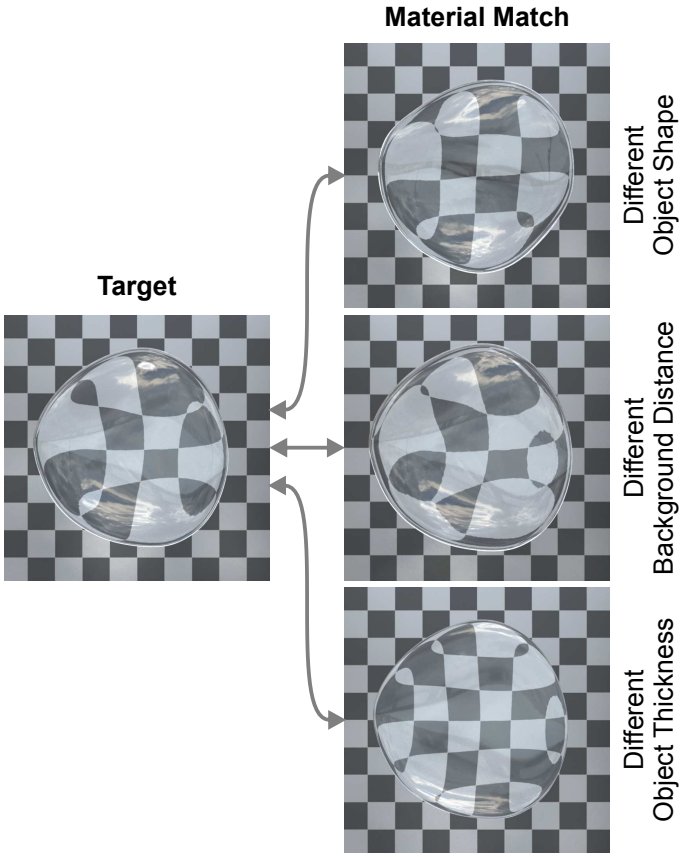

Figure 2.3: Matching situations in which standard and test object differ in either the object shape, the background distance, or the object thickness. If subjects were able to successfully match refractive material properties of transparent objects, the displayed objects should represent the best match (despite the clearly different background distortions), because they all have the same refractive index. Fleming, Jäkel, and Maloney (2011b) implemented only two of these types of context changes, namely changes in background distance and object thickness.

Fleming et al. (2011b) found that the refractive indices, adjusted by the subjects, correlated highly with those simulated in the standard objects. From this the authors conclude that the optical distortion of the background is used by the visual system to infer refraction-related material properties of transparent objects. Accordingly, it is assumed that the subjects' settings reflect abstract matches between estimated refractive indices that are part of the mental representation of transparent materials.

However, if one takes into account that the optical distortion of such objects depends, in a highly complex way, on specific values of several context factors that are – at least in the situation that was simulated in the experiments – all unknown to the subjects, this interpretation seems not very plausible. In this work, we will deal in detail with the interpretation proposed by Fleming et al. (2011b) and discuss alternative interpretations.

We will start by analyzing (a) which factors influence how the background texture is distorted, (b) which factors influence the detectability of an existing distortion, and (c) which context information and processing steps would be needed to use distortion as a cue for refraction-related material properties.

Our analysis starts from the well-known fact that optical distortions caused by thick, transparent objects are the result of a complex interplay of refraction-related material properties, object shape, parameters of the surrounding scene, and the viewing conditions. It seems nearly impossible to disentangle this interplay in order to determine the contribution of a single factor, for instance that of the refractive index, without exactly knowing the values of all other factors. This ambiguity can explicitly be demonstrated for the type of static situations that was used by Fleming et al. (2011b). We show that in these scenes objects of different thicknesses and refractive properties can cause indistinguishable background distortions.

Refraction does not only influence background distortion, but also the amount of specular reflection at the surface of the object, a fact that Fleming et al. (2011b) mentioned, but did not focus on. A brief analysis of specular reflection will reveal that it is, just like background distortion, not only influenced by the refractive index but also by several context factors. Thus, also in this case, it is unclear how it could be used as a cue for material-related properties.

We will also discuss an alternative interpretation of the findings of Fleming et al. (2011b). Our central assumption is that the subjects did not compare internal estimates of material parameters but instead performed simple matches based directly on the similarity of image attributes. Such similarities on the image level can refer to properties of the background distortion and/or characteristics of the specular reflection component. In the following, we will call these image-based matches "similarity matches" to distinguish them from abstract matches on a representational level as they are assumed by Fleming et al. (2011b).

We will present results from matching experiments similar to the ones used by Fleming et al. (2011b), in which we determined how these two kinds of image attributes influence the subject's matching behavior. The results suggest that the subject's settings reflect a compromise

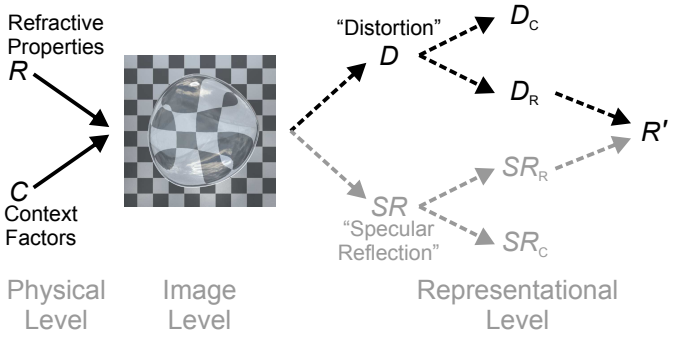

Figure 2.4: Illustration of the steps required for estimating refractive material properties of transparent objects. The black lines explicate the hypothesis proposed by Fleming, Jäkel, and Maloney (2011b) with respect to background distortions; the gray lines extend the principle to specular reflections. The starting point is a physical scene containing a transparent object. Besides the material-related refractive properties (R) of the object, it is determined by various context factors (C; e.g., the object shape and thickness, the illumination, the background distance, etc.). The visual system would first have to extract the components of the background distortion (D) and the specular reflection (SR) from the proximal stimulus. Afterwards, a separation of the two image attributes in refractive- and thus material-related components (D_R and SR_R) as well as context-related components (D_C and SR_C) would be required. The material-related components of both image attributes would then have to be used to generate a representation of refraction-related material properties (R').

between a similarity match based solely on background distortions and a similarity match based solely on specular reflections. The relative weight with which the two image attributes influence the matching behavior appears to depend on their particular salience.

2.2 Background distortion and specular reflection as potential cues for estimating the refractive index

If one assumes that the subjects' settings in the experiment of Fleming et al. (2011b) reflect a comparison between internal estimates of the refractive indices of the standard and test object, then this implies that relatively precise estimates are possible on the basis of the information provided in the stimulus. With respect to background distortions, this would require the visual system to first determine the background distortion from the proximal stimulus and then to isolate the part that is refraction-dependent (i.e., material-related) from the total distortion (see Figure 2.4).

There are different ways of how the total distortion can be determined. One obvious possibility is to compare the distorted region of the proximal stimulus with an undistorted initial state. Since the undistorted state of the region behind the transparent object is naturally not directly available, it has to be actively generated and represented in an appropriate way. To this end, the visual system might relate to an undistorted part of the background, provided

there is one in the proximal stimulus. In dynamic scenes (e.g., with changing object positions and orientations), in which the background is temporally visible in plain view, the undistorted state could be represented in memory. But even if undistorted regions are not available in the stimulus, the visual system might resort to a set of internalized assumptions about general regularities of the environment or simplicity rules as they are suggested by Gestalt theories to infer the undistorted reference texture. If, for instance, the distortion patterns in Figure 2.3 would be presented without the undistorted surround, then regularity considerations might nevertheless suggest that the undistorted texture is a checkerboard. Furthermore, if a fixed background is seen under slightly varying conditions, then the estimate of the undistorted background may improve over time due to an accumulation of partial information.

The extent to which the visual system can determine the refraction properties of the filter material from the distortion in the proximal stimulus is, in general, limited by the regularity and density of the background texture. It is more difficult to decide with irregular backgrounds than with regular ones, whether a proximal pattern originates from optical distortions elicited by a transparent object, or whether it is already contained in the undistorted background texture. Background texture density poses another limit, because an optical distortion can only be detected to the extent that it leads to noticeable texture changes in the proximal stimulus. Here, the worst case is a homogeneous background, where the proximal image remains completely unchanged under arbitrary optical distortions.

To derive refraction-related material properties from the distortion information generated in the second step, it would be necessary to isolate the refraction-dependent (i.e., material-related) part of the background distortion from the part caused by context factors. Context factors are, amongst others, the shape of the object and its position and orientation relative to the background and the observer. To be able to decompose the background distortion into a refraction-related and a context-related part, it would not only be necessary to possess exact information about the actual values of the parameters that characterize each context factor but also knowledge about the laws that describe their joint influence on the background distortion. It is not very plausible that all these requirements are fulfilled, especially for static stimuli such as those used by Fleming et al. (2011b).

2.2.1 Compensation of thickness and refractive index

The ambiguity of background distortions with respect to the refractive properties of the transparent objects that caused them becomes even more apparent if one considers the influence of object thickness. An interesting observation made in our simulations of transparent objects was that the optical distortion can be perceptually indistinguishable for objects that differ in thickness and refractive index. Figure 2.5 shows an exemplary pair of objects that produce virtually indistinguishable background distortions (an effect abbreviated below as “R-T com-

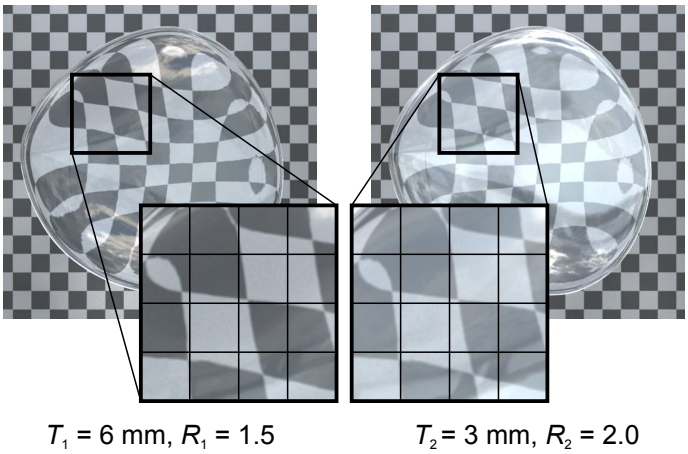

Figure 2.5: Two objects that differ in thickness and refractive index but lead to virtually identical background distortions. The left object has a maximum thickness T1 of 6 mm and a refractive index R_1 of 1.5 mm. The right object is only half as thick ($T_2 = 3$ mm). If its refractive index R_2 is chosen to be 2.0, then the background distortion caused by both objects is virtually identical. Obviously, this does not hold for the degrees of specular reflection at the objects surface, which is, due to the larger refractive index, much higher on the right side.

pensation"). To examine the regularities of this compensation in a more systematic manner, we analyzed the simple case of objects with parallel surfaces (see Figure 2.6a). To simplify things further, we first considered the optical displacement of a single ray. Figure 2.6b shows for a specific angle of incidence ($\theta = 5°$) and a reference object with refractive index R_1 and thickness T_1, how the refractive index R_2 of a second object with thickness ratio T_2/T_1, must be chosen to achieve the same optical displacement of the ray in both objects (see Appendix 2.A for a detailed derivation). The optical displacement D does not only depend on the object thickness T and the refractive index R, but also on the angle of incidence θ, with which a ray impinges upon the surface of the object. Thus, depending on the angle of incidence, different combinations of refractive indices and object thicknesses cause the same optical displacements of the corresponding ray. In general, if an observer looks at a transparent object, the angle of incidence may be different for each point on its surface (see Figure 2.6c). The higher the viewing distance is and the smaller the object is, the smaller the differences in the angle of incidence are. For the complete optical distortion, which can be regarded as a combination of optical displacements of single rays, no simple solution exists, that is, a solution with a compensating refractive index that is constant across space. Figure 2.6b shows, however, that the influence of the angle of incidence θ onto the refractive index can be neglected for a large range of T_2/T_1.

Different combinations of refractive indices and object thicknesses that cause virtually indistinguishable background distortions can also be found for more complex scenes, like the one shown in Figure 2.1c. We used numerical methods to analyze the situation illustrated in

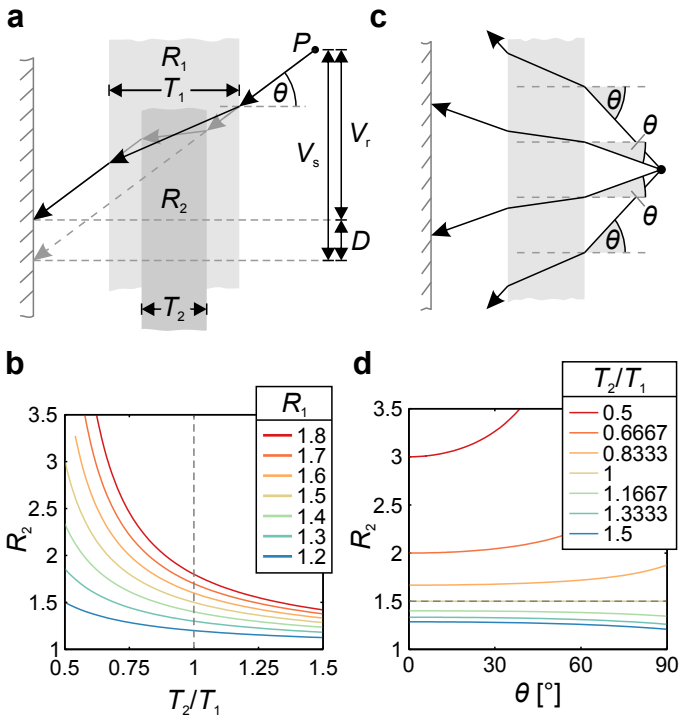

Figure 2.6: The influences of object thickness and refractive index on background distortions partly compensate each other. See Appendix 2.A for details. **(a)** Situation in which two objects with parallel surfaces (dark gray and light gray area) with different thickness ($T_1 > T_2$) and different refractive index ($R_1 < R_2$) cause the same optical displacement D, because here the larger refractive index R_2 compensates completely for the effect of the smaller thickness T_2. **(b)** For the situation illustrated in Figure 2.6a and an angle of incidence of $\theta = 5°$, the graphs show how the refractive index R_2 of an object with a thickness ratio of T_2/T_1 has to be chosen to create the same optical displacement as a second object with refractive index R_1 and thickness T_1. **(c)** The compensation of the effects of object thickness and refractive index depends on the angle of incidence θ. **(d)** The graphs show how the refractive index R_2 of an object with a thickness ratio of T_2/T_1 has to be chosen to create the same optical displacement as a second object with refractive index R_1 (shown here for $R_1 = 1.5$) and thickness T_1 depending on the angle of incidence. For many cases, these graphs are relatively flat, so that a constant value for R_2 represents a good approximation. In these cases, an approximate compensation of object thickness and refractive properties (R-T compensation) can be assumed. However, for large θ and thickness ratios less than 1, there are cases where either no solution for R_2 exists or where the values of R_2 would be higher than those typically found in nature.

Figure 2.5: Figure 2.7a shows (averaged across all angles of incidence) which refractive index R_2 of an object with thickness T_2 causes an optical distortion that is virtually indistinguishable to the one caused by a second object with refractive index R_1 and thickness T_1. The results are again plotted against the thickness ratio T_2/T_1. A comparison with Figure 2.6b reveals that the graphs are similar to those obtained for the simplified situation. Figure 2.7b shows that the residual differences between the distortions that were considered as being indistinguishable are indeed negligible for the range of values used in Figure 2.7. In summary, the analysis reveals that different combinations of thicknesses and refractive indices may cause virtually indistinguishable background distortions. In particular, this is the case for the parameter range used in Fleming et al.'s (2011b) experimental paradigm ($2/3 < T_2/T_1 < 2$ and $1.1 < R_1 < 2.0$.

2.2.2 Specular reflection as a cue

Fleming et al. (2011b) did not consider in detail that the refractive index of transparent objects also influences how much of the incident light is specularly reflected from their surfaces. The higher the refractive index of a transparent object is, the stronger these specular reflections are (see Figure 2.2). Additionally, the higher the angle of incidence is, the more light is reflected (Fresnel effect).

In general, using specular reflections as a cue for the refractive properties of transparent objects leads to similar problems as the use of background distortions, because specular reflections also vary with several parameters of the scene that are not related to material properties. They depend, for instance, on the illumination or the viewpoint. Thus, like background distortions, specular reflections are also not uniquely related to the refractive index. To estimate refraction-related material properties from specular reflections, the visual system would first need to generate information about the specular reflection from the proximal stimulus. Afterwards, this information would have to be decomposed in a refraction-dependent (i.e., material-related) and a context-dependent part. Like in the case of background distortions, this decomposition would require exact knowledge about all context factors and their specific influence on the specular reflection (see Figure 2.4).

2.2.3 Discussion

The previous analysis revealed that the background distortion elicited by a transparent object is not uniquely related to refraction-related material properties, but instead depends on many additional context factors. Moreover, objects that differ in refractive index and thickness can cause virtually indistinguishable distortions. In order to use the background distortion to estimate refraction-related material properties, distortion information has to be extracted from the proximal stimulus. The detectability of the distortion strongly depends on context factors like the characteristics of the background itself. Furthermore, the total distortion

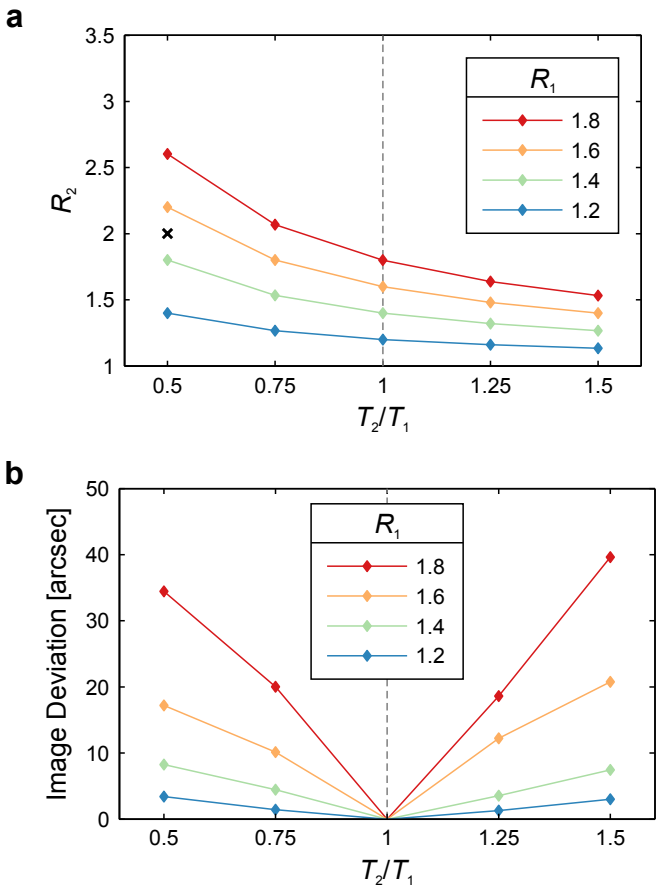

Figure 2.7: Compensation of refractive index and object thickness (R-T compensation) for the situation illustrated in Figure 2.5. **(a)** The graphs show how the refractive index R_2 of an object with thickness T_2 has to be chosen to achieve a background distortion that is virtually indistinguishable from the one caused by a second object with refractive index R_1 and thickness T_1. The values are plotted against the thickness ratio T_2/T_1 and represent averages across all angles of incidence. Because no simple analytic solutions exist for these situations, ray paths were simulated explicitly. The black cross represents the case shown in Figure 2.5. **(b)** The graphs show the mean difference between corresponding points in the distorted images obtained with the two objects, if the refractive index R_2 is chosen as illustrated in Figure 2.7a. The largest average deviation is less than $40''$ of viewing angle. Thus, the corresponding background distortions are virtually indistinguishable and the influence of the angle of incidence θ is negligible. In this example, $0.01°$ viewing angle corresponds to only $0.14\,\%$ of the object diameter in the image or to 0.26 px in an image of size 370×370 px.

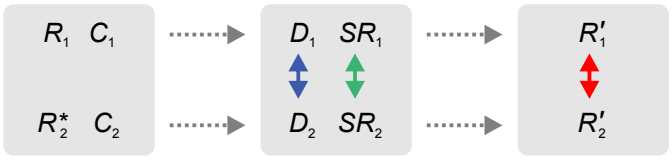

Figure 2.8: Schematic depiction of different possible matching behaviors in situations like the one used by Fleming, Jäkel, and Maloney (2011b) in which the material of a given object with fixed refractive index R_1 presented in context C_1 must be matched in a second object presented in context C_2, by adjusting its refractive index R_2^*. Fleming, Jäkel, and Maloney (2011b) proposed that the settings of R_2^* were performed by comparing internal representations of refractive indices that were estimated from the proximal stimulus (R' match, red arrow). Alternatively, subjects might have performed simple similarity matches based directly on image attributes related to either background distortions (D match, blue arrow) or specular reflections (SR match, green arrow) or a combination of these.

information would have to be decomposed into a refraction-dependent (i.e., material-related) and a context-dependent part. This would require extensive knowledge about the actual values of all relevant context factors and how they influence background distortion. Putting all the facts together, it thus seems highly questionable whether the background distortions can be used, at all, to infer refraction-related material properties of transparent objects. In particular, this speaks against the interpretation that subjects compare internal representations of estimated refraction indices in matching tasks like the one of Fleming et al. (2011b).

An alternative explanation is that subjects perform simple matches based directly on image attributes available in the proximal stimulus (see Figure 2.8). This would be a reasonable strategy in the given experimental situation, because both objects are visible at the same time and can be compared directly. This explanation seems to be compatible with the results of Fleming et al. (2011b): Even though the refractive indices chosen by the subjects for the test objects correlate with those of the standard objects, they nevertheless show systematic deviations. The specific form of these deviations suggests a matching behavior that was based on similarity matches of background distortions (abbreviated below as “D match”). Moreover, in their exemplary illustration of the mean settings made by their subjects, a highly similar background distortion in standard and test object can be observed (Fleming et al., 2011b, Figure 3d). Because background distortions are not the only image attribute available in the stimulus, such similarity matches could just as well refer to specular reflections (abbreviated below as “SR match”), or the subjects’ settings might represent a compromise between both kinds of similarity matches.

It is presently unclear on which image criteria such similarity matches might be based. In principle, a large number of criteria is possible. For example, summary measures like the maximal brightness of the specular reflection or the average element size of the distorted background might be used. Alternatively, more complex scene statistics like the brightness

histogram of specular reflections or histograms of texture element sizes could be used. Such similarity matches can be performed even if exact proximal identity matches are not possible.

It is important to point out that the results of similarity matches based on specular reflections (SR match) can easily be misjudged as resulting from successful abstract matches of estimated refraction indices (R' match). Because the specular reflections are hardly influenced by the context factors varied in the experimental paradigm used by Fleming et al. (2011b), SR matches would accidentally also lead to good matches of the refractive index. This means that any result that is consistent with an R' match, can be explained more economically by a simple SR match.

A further interesting point is that similarity matches, like the ones outlined above, are still possible even if the stimuli do not elicit an impression of transparency. This clearly indicates that an explanation of the subjects' settings in terms of similarity matches is of a fundamentally different nature than the one proposed by Fleming et al. (2011b) themselves, because the latter relates to estimates of refraction indices and therefore only appears reasonable if a transparency-specific mechanism of the visual system is assumed. The question on which basis the subjects actually made their settings is therefore of great theoretical importance.

2.3 Experiment 1

In order to investigate which strategy subjects use if they are asked to match the perceived material properties of two transparent objects, an experimental situation is required in which the use of different matching strategies leads to distinguishable results. The experimental paradigm of Fleming et al. (2011b), in which the material of transparent objects of different thickness must be matched, seems well-suited for this purpose.

If subjects were able to successfully perform an abstract match of refraction-related material properties (R' match), then the refractive indices set in the adjustable test objects should be identical to those used in the corresponding standard objects. Contrary to this expectation, systematic deviations from the given refractive indices were observed in the results of Fleming et al. (2011b). In principle, these deviations could at least in part be due to contradictory information in their stimuli. To reduce the likelihood of such negative influences in our study, we tried to further optimize the realism of the stimuli. Due to these changes, our stimuli differ slightly from those used in Fleming et al. (2011b). All perspective parameters of our stimuli were chosen in correspondence with the actual settings and physical apparatus used in the experiment. Thus, with respect to geometry, the stimuli appeared virtually the same as a real instance of that scene would have appeared. Furthermore, we used complex natural environmental lighting instead of a few simple localized light sources, because this reduces high-contrast highlights that are difficult to handle in tonemapping. As a result, the salience of the specular reflections may be slightly higher in our stimuli.

If subjects would perform similarity matches based on background distortions (D match), then the adjusted refractive indices would differ from the ones given (as already mentioned, simple D matches lead in nontrivial matching situations as those realized by Fleming et al. (2011b) to nonveridical refractive indices). The approximate R-T compensation discussed above can be used to predict for standard and test objects of different thickness the settings of the refractive index that should result, if the subjects actually perform similarity matches of background distortions: It would be the refractive index that leads to virtually indistinguishable background distortions in standard and test, because this should always represent the best match, no matter which (potentially abstract) criteria are actually used in a similarity match of background distortions.

Predictions for similarity matches based on the specular reflection (SR match) were gained empirically by asking subjects to match objects with isolated specular reflections. For this purpose, we used homogeneous background textures that cannot map any background distortions. Thus, the specular reflection remains the only base for the matches.

2.3.1 Stimuli

The stimuli were computer-generated images created with the Mitsuba renderer (Jakob, 2013). The stimuli were created to closely resemble the ones used by Fleming et al. (2011b) to ensure comparability. The stimulus images showed a thick transparent object in front of a background board that was located inside a box with front and top sides open (see Figure 2.9). All scene elements were defined in real-world coordinates relative to a virtual projection plane, which represented the surface of the experimental screen. The shape of the transparent object was based on an icosahedron that was subdivided seven times. The resulting icosphere was deformed to a slightly warped ellipsoid with the computer graphics software Blender (Blender Foundation, 2013) by applying various shape modifiers to its mesh. The object size was about 50×50 mm (width, height). Depending on the experimental condition, the thickness of the standard object varied in four steps ($T_S \in \{3, 4.5, 7.5, 9\,\text{mm}\}$), while the thickness of the test object remained constant ($T_T = 6\,\text{mm}$). The thickness was manipulated by applying scaling factors to the object mesh. The object was made of light transmitting material ("dielectric") without any internal absorption. The refractive indices of the standard stimuli varied according to the experimental condition in seven steps ($R_S \in \{1.20, 1.30, 1.40, 1.50, 1.60, 1.70, 1.80\}$). The refractive index of the test stimuli were adjusted by the subjects ($R_S \in \{1.010, 1.015, \ldots, 2.495, 2.500\}$, 299 steps). The refractive index of the medium surrounding the transparent object was set to 1. The transparent object was located at the center of the virtual projection plane. A textured 80×80 mm background board was placed behind the transparent object at a distance of 60 mm. The background textures were random Voronoi patterns created with MATLAB (Mathworks, Inc., Natick, MA) that resembled the background textures used by Fleming et al. (2011b). The

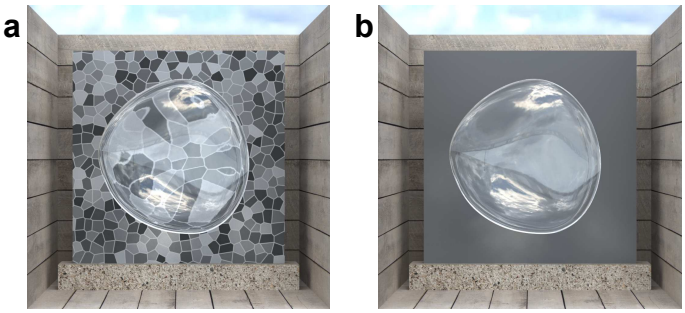

Figure 2.9: Illustration of the stimuli used in the first experiment. **(a)** In half of the trials we used a random Voronoi pattern as background texture that resembled the background textures used by Fleming, Jäkel, and Maloney (2011b). **(b)** To gain predictions for similarity matches based solely on the specular reflection (SR match), we used a homogeneous background texture in the other half of the trials.

individual faces of the pattern used in the rendering were separated by seams with a width of 0.32 mm and a color of RGB(.706, .706, .706). The colors of the faces were uniformly distributed between RGB(.294, .294, .294) and RGB(.686, .686, .686). Additionally, we used a homogeneous background texture (RGB(.490, .490, .490)) to isolate the specular reflection component. The sole illumination was provided by an infinitely distant high dynamic range sphere emitter, containing a natural daylight outdoor scene with a partly cloudy sky. The camera settings (location and field of view) were chosen to correspond to the actual experimental setup. Thus, the stimuli appeared in exactly the same way as a corresponding real scene, and there were virtually no perspective distortions. Stimuli were rendered as 16 bit/channel high dynamic range images ("extended volumetric path tracer"; "low discrepancy sampler" with 512 samples/px; "Gaussian reconstruction filter" with $SD = 0.5$) and subsequently tonemapped to 8 bit/channel low dynamic range images ($\gamma = 1.6$; exposure = 1.4). The final image size was 370×370 px which corresponded to 100×100 mm on the screen.

2.3.2 Procedure

In each trial, two stimuli were presented simultaneously on an Eizo ColorEdge CG243W LCD screen (display area 518.4×324 mm; resolution 1920×1200 px; color depth 8 bit/channel; 3.704 px/mm; Eizo Corporation, Hakusan, Japan) in a darkened room. The fixed standard stimulus was located at the top of the screen, the adjustable test stimulus at the bottom. The viewing distance was 400 mm. The subject's task was to adjust the refractive index of the test object until it appeared to be made of the same material as the fixed standard object. The settings were made with the arrow keys on a standard computer keyboard.

Each subject performed 168 trials in randomized order. In 84 of them, the Voronoi texture was used ($7R_S \times 4T_S \times 3$) and in the remaining 84 the homogeneous texture ($7R_S \times 4T_S \times 3$).

2.3.3 Subjects

Six subjects, three of them female, participated in the experiment. Their ages ranged from 20 to 35. All subjects were naive as to the purpose of the experiment. They reported normal or corrected-to-normal visual acuity, and showed no color vision deficiency, as tested by Ishihara plates (Ishihara, 1969).

2.3.4 Results

Figure 2.10 shows the results of the first experiment. For all four object-thickness ratios (T_T/T_S) and all refractive indices of the fixed standard (R_S), the mean setting of the refractive indices of the test object lies between the predictions for a similarity match of background distortions (D match) and a similarity match of specular reflections (SR match). The settings of the refractive indices R_T averaged across all object thickness ratios T_T/T_S lie closer to the predictions for similarity matches of specular reflections (SR match) if the refractive indices R_S of the fixed standard is high (see black curve in Figure 2.11a) and tends towards the predictions for similarity matches of background distortions (D match) for low R_S. For an object thickness ratio of $T_T/T_S = 100/150$ and a standard object with a refractive index of $R_S = 1.5$, Figure 2.11b illustrates exemplarily how the test object would have looked like for a D match, a SR match, and for the subjects' mean setting for the refractive index.

Our findings closely resemble the corresponding results reported in Fleming et al. (2011b), although this is somewhat hidden by the different way in which our results are presented: The refractive indices adjusted by the subjects correlate highly with the refractive indices of the standard objects, but show systematic deviations. This suggests that both experiments probe the same processes, despite the slight differences between the stimuli used in both experiments.

The settings of two subjects differ substantially from the results presented so far and are therefore considered separately. Regardless of the object thickness ratio (T_T/T_S) or the refractive index of the standard object (R_S), their settings always resemble the ones expected for a similarity match based solely on background distortions (D match; Figure 2.12).

2.3.5 Discussion

In our experiment, we used matching situations in which different matching strategies would lead to clearly discriminable settings. The range of the adjustable parameter of the test stimulus always enclosed the value predicted for similarity matches based solely on the background distortions (D match) and the value predicted for similarity matches based solely on the specular reflections (SR match). By comparing the actual settings with these two predictions, it is possible to gain information about the relative influence of both image criteria on the

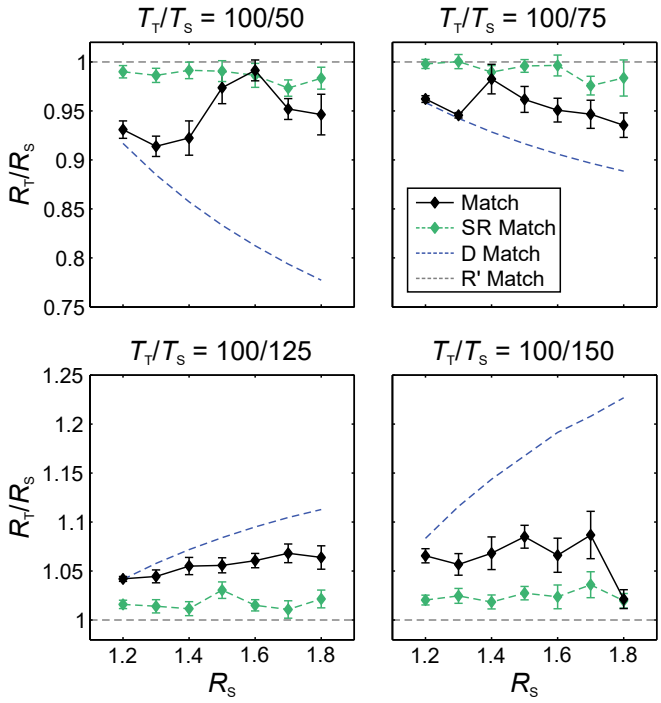

Figure 2.10: Results of the first experiment. Each plot corresponds to one object thickness ratio T_T/T_S. The black solid lines represent the mean setting of the refractive indices of the test object. The dashed blue lines represent the predictions for similarity matches based solely on the background distortions (D match), which were gained by simulations. The dashed green lines represent the predictions for similarity matches based solely on specular reflections (SR match), which were gained by matches with isolated specular reflections. The dashed gray lines represent the settings expected for abstract matches of estimated refractive indices (R' match). All values are depicted as ratios of the refractive index of the test object to the refractive index of the standard object (R_T/R_S, $\pm SEM$).

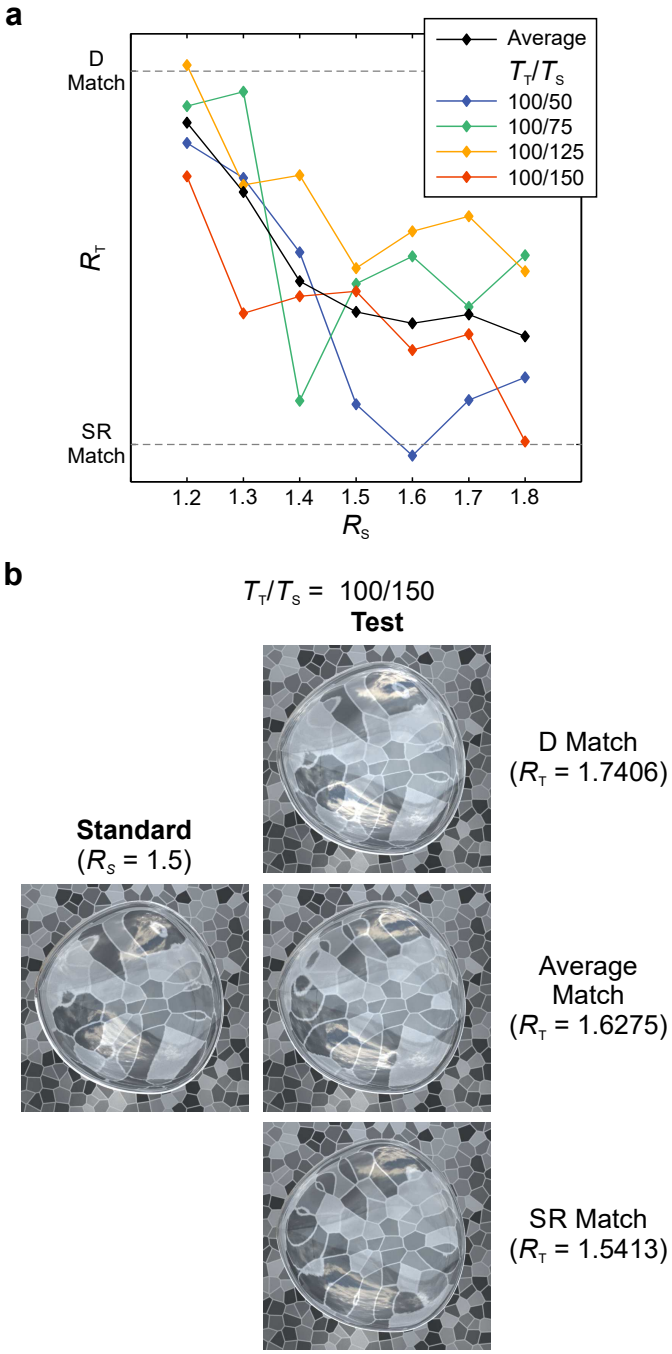

Figure 2.11: The settings for the refractive indices lie between the predictions for similarity matches of background distortions (D match) and similarity matches of specular reflections (SR match). **(a)** Relative position of the refractive indices of the test object (R_T) in the interval between the respective predictions of D matches and SR matches for all four thickness ratios (T_T/T_S) as a function of the refractive index of the standard object (R_S). The black line represents the average setting across all four thickness ratios. **(b)** Example of a fixed standard object (left column, $R_S = 1.5$) and a test object (right column, $T_T/T_S = 100/150$) with different settings for the refractive index R_T. The topmost object shows a similarity match based solely on background distortions (D match, gained by simulations using the R-T compensation); the lowermost object shows a similarity match of specular reflections (SR match, gained empirically). The refractive index of the center object corresponds to the subjects' mean setting. The images show only a part of the complete stimulus illustrated in Figure 2.9a.

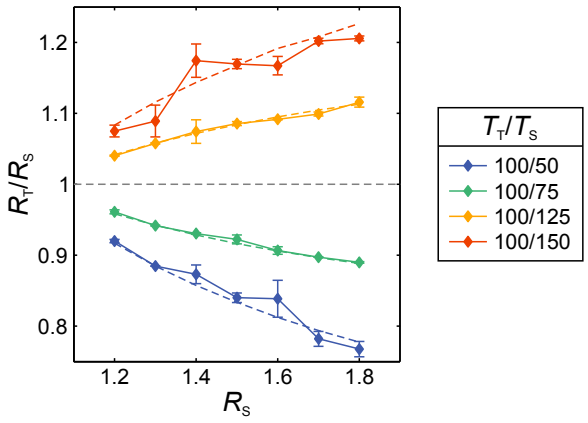

Figure 2.12: The settings of two of the subjects deviate systematically from those of the other subjects. Their adjusted refractive indices are very similar to the predictions made for similarity matches of background distortions (D match). The colored solid lines represent the ratios of the refractive indices of the adjustable test to the fixed standard (R_T/R_S, $\pm SEM$) for a particular object thickness ratio (T_T/T_S). The dashed lines represent the D match prediction for each corresponding object thickness ratio.

matching behavior. Our results suggest that the matches reflect a compromise of similarity matches based on background distortions and similarity matches based on specular reflections. The relative influence of the two image attributes seems to depend on their salience. With lower refractive indices of the standard object and thus weak specular reflections, the subjects tended to base their setting mainly on background distortions. With higher refractive indices of the standard object and correspondingly stronger specular reflections, they tend to rely more on specular reflections.

Apparently, two of our subjects always performed similarity matches of background distortions and completely ignored specular reflections. The fact that different subjects used different matching strategies suggests that subjects deliberately choose a matching criterion. This suggests that also the compromise between similarity matches of background distortions and similarity matches of specular reflections that was found for the other subjects might mainly be the result of a conscious decision and not the result of unconscious cue integration processes in visual perception.

2.4 Experiment 2

Our alternative interpretation of the results of Fleming et al. (2011b) is that their subjects did not compare internal estimates of refractive indices, but instead performed simple similarity matches of image attributes in the proximal stimulus. It is presently unclear to which criteria such similarity matches actually refer. It might be possible that summary measures, such as the average size of texture elements or the maximum brightness of the surface reflection,

play an important role. In any case, it seems clear that such similarity matches must rely on abstract similarity criteria and not on local proximal identity. To test the hypothesis that such similarity matches are not limited to matching situations in which exact proximal identity of an image attribute could be achieved (as was the case for the background distortion and, to a certain extent, also for the specular reflection in the first experiment), we conducted a matching experiment in which the object shape differed between the standard and test object. Thus, in contrast to the first experiment, neither the background distortion nor the specular reflection could be made proximal identical. Both the background distortions and the specular reflections were matched in isolation from each other.

If the results of this experiment are comparable to the ones of the first experiment, then the hypothesis that subjects perform abstract similarity matches based on the background distortion and the specular reflection is supported.

2.4.1 Stimuli

The stimuli were largely the same as in the first experiment. However, different object shapes were used for the standard stimuli by applying different shape modifiers. As in the first experiment, the refractive index of the standard object was varied in seven steps ($R_S \in \{1.20, 1.30, 1.40, 1.50, 1.60, 1.70, 1.80\}$). To isolate information given by the specular reflection component, we used the same homogeneous background texture as in the first experiment (RGB(.490, .490, .490)). In this condition, standard and test objects did not differ in any parameter, except in shape. To isolate information given by background distortions, we rendered stimuli without the specular reflection component by choosing a corresponding option in the Mitsuba renderer (Jakob, 2013). Because higher refractive indices lead to a higher degree of specular reflection, the deactivation of the specular reflection component darkened the image area of the object. To compensate for this, we adjusted the brightness of the object area accordingly. In this condition, the thickness of the standard object was varied in four steps ($T_S \in \{3, 4.5, 7.5, 9\,\text{mm}\}$). The object thickness was the same for all test stimuli ($T_T = 6\,\text{mm}$), while their refractive properties varied according to the inputs made by the subject ($R_T \in \{1.010, 1.015, \ldots, 2.495, 2.500\}$, 299 steps).

2.4.2 Procedure

The procedure was the same as in the first experiment. Every subject was presented with 105 trials in randomized order. Twenty-one of the trials had an isolated specular reflection component ($7R_S \times 3$), and 84 had an isolated background distortion ($7R_S \times 4T_S \times 3$).

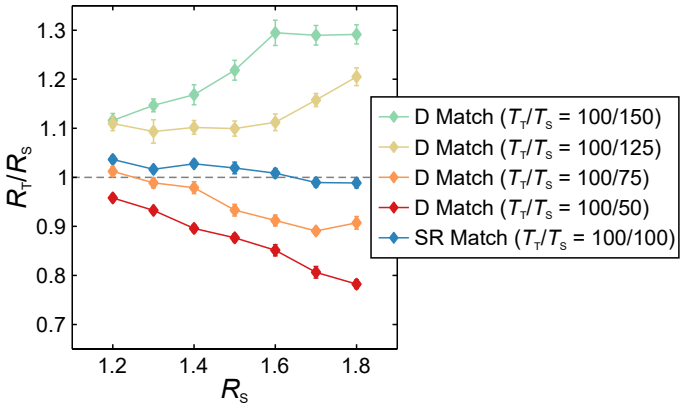

Figure 2.13: Results of the second experiment. Ratios of the mean refractive index settings to the refractive index of the corresponding standard object (R_T/R_S, $\pm SEM$) for matches with isolated distortion (D match) under four object thickness ratios (T_T/T_S), and for matches with isolated specular reflection (SR match).

2.4.3 Subjects

The subjects were the same as in the first experiment.

2.4.4 Results

Figure 2.13 shows the mean settings of the refractive indices of the test object for the matches with isolated background distortions (D match) and for matches with isolated specular reflections (SR match). For situations with isolated background distortions, the settings are similar to the predictions for similarity matches of background distortions derived in the first experiment. Hence, these settings are also similar to those of the two subjects of the first experiment, who based their matches solely on background distortions.

The matches in situations with isolated specular reflections are similar to the predictions for similarity matches of specular reflection, which were gained in the first experiment in a similar way (but with the same object shapes in the standard and test stimulus).

2.4.5 Discussion

The data show that subjects are able to perform reliable similarity matches based on background distortion and specular reflection, even if the objects in the standard and match stimulus differ in shape so that neither the background distortion nor the specular reflection could be matched directly (i.e., by proximal identity). For both kinds of similarity matches, results similar to the ones of the first experiment were found. This suggests that subjects refer to abstract similarity measures, such as, for example, the average size of texture elements in

the distorted background area or the maximal brightness of the specular reflection, and not to proximal identity.

2.5 General Discussion

Fleming et al. (2011b) tested the hypothesis that background distortions are used to perceive refraction-related material properties of thick transparent objects. In their experiments, they asked subjects to match the refractive indices of two transparent objects. From the fact that the refractive indices set by the subjects correlated highly with the given ones, they conclude that background distortions are indeed used to perceive material properties of transparent objects. They interpret the subjects' settings as the result of abstract matches of refractive properties – that is, they assume that the subjects compared internal representations of refractive indices that were estimated from information given in the stimulus.

Background distortions and specular reflections are two image attributes that depend on the refractive index, a material property of light transmitting objects, and may thus be regarded as potential cues that can be used to infer refraction-related properties from the proximal stimulus. However, an analysis of these image attributes reveals that they do not only depend on the refractive index but that they are also influenced by several unknown context factors. Thus, these image attributes can only serve as cues for the refractive index, if it would be possible to decompose them into a refraction-dependent (i.e., material-related) part and a context-dependent part. We argue that this assumption is not very plausible, especially if static stimuli are used as in Fleming et al. (2011b).

We tested in two experiments the alternative explanation that subjects do not perform abstract matches of estimated refractive properties, but simple similarity matches of image attributes. In line with this hypothesis, our results indicate that the matches reflect a compromise between a similarity match of background distortions (D match) and a similarity match of specular reflections (SR match). The weight that is given to each image attribute in the match seems to partly depend on its salience. Our results indicate that specular reflections tend to be the dominant factor as soon as they become clearly noticeable. Background distortion dominated the settings only for low values of the refractive index of the standard object (and thus weak specular reflections). However, these findings should be interpreted with caution, because the additional finding that the relative contribution of the two image attributes differed considerably between subjects suggests that the compromise is mainly of a cognitive nature.

It is presently unclear which image-based criteria are used in such similarity matches. At any rate, the results of our second experiment, in which subjects matched perceived materials across objects of different shape, indicate that subjects use abstract similarity criteria in such matches and do not refer to local identity. This was found both for isolated background

distortions and isolated specular reflections. It remains a goal of future work to find out to which criteria similarity matches of the background distortion or the specular reflection actually refer.

Our theoretical analysis revealed that it is highly implausible that background distortions caused by thick transparent objects can be used to gain an estimate of a refractive index that serves as an internal representation of refraction-related material properties. This judgment is primarily based on computational arguments, but there are several additional arguments that speak against the idea that the visual system tries to gain an estimate of the refractive index as part of an internal representation of light transmitting materials.

A first point is that most of the light-transmitting materials in our environment are either water ($R = 1.33$) or (acrylic) glass ($R \approx 1.50$ up to 1.70). Other exemplars are much more rare, like diamonds ($R = 2.42$). It is therefore not to be expected that a metrical representation of refractive properties would substantially improve the functional linking of the visual system to the world surrounding us. A second point is that the matching accuracy demonstrated in the experiment of Fleming et al. (2011b) and in the experiments reported here is relatively low compared to the small range of relevant refractive indices found in the natural environment. Thus, in most real world cases one could probably reduce the estimation error considerably, if one would replace the estimate from the proximal stimulus with one of the values 1.33 or 1.5. A further point is that a representation of the refractive index would only be useful if its *absolute* value were exactly known. However, such absolute representations are at least very rare in vision; almost all representations are of a relational nature.

Clearly, this does not mean that background distortions and specular reflections of transparent objects are ignored by the visual system. To the contrary, it seems highly probable that this information is used in many different ways in transparency perception. For example, it seems obvious that specular reflections play an important role for shape perception, just as in the case of opaque objects (see e.g. Blake & Bülthoff, 1990, 1991; Fleming, Torralba, & Adelson, 2004; Norman, Todd, & Orban, 2004b). Specular reflections of transparent objects are more complex because reflections at different surfaces (e.g., at the front and at the inner back of the object) can be superimposed in the retinal image. At present, it remains unclear if this higher complexity of specular reflections enhances the perception of the object shape (and maybe even allows perceiving the shape of the back surface, which is not directly visible) or if it makes shape perception more difficult.

With respect to background distortions, it appears plausible that they act as an unspecific cue, which plays an important role in identifying objects as being transparent, especially in situations where no other cues for transparency (e.g., color relations) exist. Background distortions can also potentially be used as a cue for the shape of the transparent object. For example, highly distorted areas could indicate significant changes in surface orientation. The use of background distortions as a cue for object shape seems more plausible than its use as a

cue for the refractive index, because relative distortion information inside the boundary of the object may suffice in the former case, but not in the latter.

Appendix

2.A Compensation of refractive index and object thickness

In addition to the parameters defined in Figure 2.6, let D_P be the distance of the object's center to the viewpoint P and D_B be the distance of the object's center to the background. To simplify calculations, the refractive index of the medium surrounding the object is assumed to be 1.

If no transparent object exists, a ray originating from the viewpoint P intersects the background at

$$V_\mathrm{S} = (D_\mathrm{B} + D_\mathrm{P}) \tan\theta. \tag{2.1}$$

If the ray is distorted by the refractive properties of the transparent object, it intersects the background at

$$V_\mathrm{r} = \frac{T \sin\theta}{\sqrt{R^2 - (\sin\theta)^2}} + (D_\mathrm{B} + D_\mathrm{P} - T) \tan\theta. \tag{2.2}$$

The difference of both intersections gives the optical displacement

$$D = \frac{T \sin\theta}{\sqrt{R^2 - (\sin\theta)^2}} - T \tan\theta. \tag{2.3}$$

The optical displacement depends on the object thickness T, the refractive index R, and the angle of incidence θ.

For two objects of different thickness (T_1 vs. T_2) and different refractive index (R_1 vs. R_2) but equal distances ($D_\mathrm{P} = D_\mathrm{B}$), the optical displacement is identical, if the thickness or the refractive index of one of the objects is appropriately chosen. For example, if the thickness of the second object is

$$T_2 = T_1 \frac{\sqrt{R_2^2 - (\sin\theta)^2}(-1 + \sec\theta\sqrt{R_1^2 - (\sin\theta)^2})}{\sqrt{R_1^2 - (\sin\theta)^2}(-1 + \sec\theta\sqrt{R_2^2 - (\sin\theta)^2})}, \tag{2.4}$$

then both objects cause the same optical displacement. The refractive index R_2 that leads to identical optical distortions can be described in a similar way by rearranging the formula.

Kapitel 3

Matching the material of transparent objects: The role of background distortions.

Dieses Kapitel wurde veröffentlicht als: Schlüter, N. & Faul, F. (2016). Matching the material of transparent objects: The role of background distortions. *i-Perception*, *7*(5), 1–24. https://doi.org/10.1177/2041669516669616 (https://creativecommons.org/licenses/by/3.0/)

Zusammenfassung

Es wurde vorgeschlagen, dass das visuelle System in der Lage ist, den Brechungsindex von dicken transparenten Objekten aus den von ihnen verursachten Hintergrundverzerrungen zu schätzen. Konkret wurde die Hypothese aufgestellt, dass dies auf Basis des Verzerrungsfelds geschieht, dessen Berechnung aus dem Eingangssignal einen Vergleich des durch das Objekt gesehenen Teils des Hintergrunds mit dem unverdeckt sichtbaren Teil erfordert. Wir testen zwei aus dieser Hypothese abgeleitete Vorhersagen: (a) Szenenvariablen, die das Verzerrungsfeld nicht verändern, zum Beispiel die Dichte der Hintergrundtextur, sollten die Einstellungen der Versuchspersonen in einer Materialabgleichsaufgabe nicht systematisch beeinflussen. (b) Die Unsicherheit der Schätzungen sollte stark ansteigen, wenn der unverdeckte Teil der Hintergrundtextur entfernt wird. Unsere Ergebnisse sind mit diesen beiden Vorhersagen nicht kompatibel, entsprechen aber vollkommen der alternativen Interpretation, dass die Versuchspersonen die Ähnlichkeit der verzerrten Hintergrundtexturen auf Bildebene maximiert haben. Zusätzliche Ergebnisse deuten darauf hin, dass die Versuchspersonen zwar das Verhältnis zwischen verzerrtem und unverzerrtem Hintergrund berücksichtigen können, sofern das experimentelle Design dies nahelegt, allerdings nur auf eine vereinfachte Weise, die nicht dazu geeignet ist, den Brechungsindex zu schätzen.

Abstract

It has been proposed that the visual system is able to estimate the refractive index of thick transparent objects from background distortions caused by them. More specifically, it was hypothesized that this is done based on a mid-level cue, the distortion field, whose computation from the input requires comparing the part of the background seen through the object with the part visible in plain view. We test two predictions derived from this hypothesis: (a) Scene variables that do not change the distortion field, for instance, the density of the background texture, should not systematically influence the subjects' settings in a material matching task. (b) The uncertainty of the estimates should increase sharply, if the part of the background texture in plain view is removed. Our results are not compatible with these two predictions but are completely in line with the alternative interpretation that the subjects maximized the similarity of the distorted background textures on the image level. Additional results indicate that subjects can take relations between the distorted and the undistorted background into account if this is encouraged by the experimental design, but they do this in a simplistic way that is inappropriate to estimate the refractive index.

3.1 Introduction

Light-transmitting materials often appear transparent. Previous investigations of perceived transparency that relate to such materials have mainly focused on their transmission properties and correspondingly considered the color and luminance relations that lead to perceptual transparency (Beck, 1978; Beck, Prazdny, & Ivry, 1984; Faul & Ekroll, 2002, 2011; Khang & Zaidi, 2002a, 2002b; Ripamonti, Westland, & Da Pos, 2004; Robilotto, Khang, & Zaidi, 2002). These investigations used highly reduced situations in which simple flat filters without refraction and scattering were viewed frontally under a homogeneous illumination. Recent progress in computer graphics made it possible to simulate more complex light-transmitting objects. This includes translucent objects like wax or milk, which due to subsurface scattering appear more or less hazy. Fleming and Bülthoff (2005) and Motoyoshi (2010) proposed specific cues that can be used to identify such objects and to characterize their properties. A further class of objects that can now be investigated are clear light-transmitting objects that are made of refractive materials but lack absorption and scattering. Light incident on the surface of a refractive material is partly reflected and partly transmitted (Figure 3.1a). The relative amount of these two parts can be predicted by Fresnel's equations and depends on the refractive index R of the object's material and the angle of incidence θ. The reflected light leaves the surface in the mirror direction (specular reflection). The transmitted light is bend at the surface by an angle that is given by Snell's law (refraction). The effects of refraction and specular reflection on the

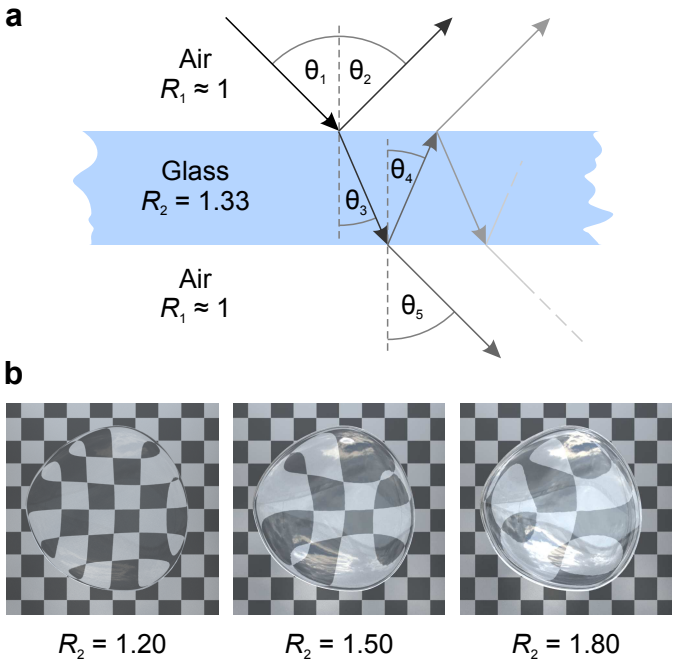

Figure 3.1: Refraction and reflection and their effects on the appearance of light-transmitting objects. **(a)** If light hits a light-transmitting object whose material has a higher refractive index (R) than that of the surround, for example, a pane of glass surrounded by air, it is partly reflected (with the angle of reflection being equal to the angle of incidence, i.e., $\theta_2 = \theta_1$) and partly refracted toward the surface normal (i.e., $\theta_3 < \theta_1$). If the refracted light hits the opposite boundary of the object, it again splits up. However, the refraction occurs in the opposite direction (i.e., $\theta_5 > \theta_3$) and higher angles of incidence can even lead to total reflection. **(b)** Effects of refraction and reflection on the retinal projection of a thick irregularly shaped light-transmitting object without absorption or subsurface scattering. The strength of specular reflection determines the brightness of the mirror image of the environment on the object's surface. The strength of refraction determines the degree of background distortion. Both the specular reflection and the amount of refraction increase with the refractive index. Note that the shape of the object is the same for all three images.

retinal projection of light-transmitting objects are especially apparent for thick, irregularly shaped objects like the ones shown in Figure 3.1b.

Due to the specular reflections, the environment is partly visible at the object's surface. This suggests that previous investigations dealing with the perception of opaque mirror surfaces that specularly reflect all incoming light (Blake & Bülthoff, 1990, 1991; Fleming, Torralba, & Adelson, 2004; Savarese, Chen, & Perona, 2004a, 2005; Savarese, Fei-Fei, & Perona, 2004b) are of some relevance. However, it is at present unclear to what extent these findings can directly be transferred to light-transmitting objects, which show a more complex reflection behavior. In these objects, the amount of reflection depends not only on the material's refractive index but also on the viewing angle: It is low at parts of the surface where the viewing direction is nearly perpendicular and approaches 1 at grazing angles. Furthermore, multiple specular reflections that originate from different surfaces, for instance, the frontal and the back plane

of a pane of glass, can be superimposed in the image. This is because specular reflections occur at each boundary of a light-transmitting object, irrespective of whether the light hits the boundary from the inside or the outside of the object.

The refraction, which is of primary importance to the present investigation, leads to distortions in the image. These are especially pronounced for thick irregularly shaped light-transmitting objects, like the blob of clear glass shown in Figure 3.1b. Two recent studies proposed that these background distortions are used by the visual system to infer object properties. Chen and Allison (2013) considered the relationship between optical distortions and object shape. Their results suggest that the pattern of distortions is actually used as a cue for shape. Fleming, Jäkel, and Maloney (2011b) referred to the regular relationship between optical distortions and the refractive index of the object's material: The larger the difference in the refractive indices at the interface between two media, the larger the distortions. They tested the hypothesis that the visual system uses optical distortions to estimate the refractive index (RI hypothesis). In their matching experiment, two thick transparent objects were presented in different static scenes, and the subjects' task was to adjust the refractive index of the test object until it appeared to be made of the same material as the standard object. They found a high correlation between the refractive indices set by the subjects and the fixed refractive indices given in the standard objects. This led the authors to conclude that "the pattern of image distortions that occurs when a textured background is visible through a refractive object provides a key source of information that the brain can use to estimate an object's intrinsic material properties" (p. 818).

Testing the RI hypothesis in this direct way is highly problematic for two reasons. A first problem is confounding variables. Physically correct stimuli as the ones used by Fleming et al. (2011b) contain not only background distortions but also specular reflections, which likewise depend on the refractive index. In Schlüter and Faul (2014), we have shown that specular reflections can substantially influence the subjects' settings in such matching tasks. It is thus difficult to discern the relative influence of background distortions and specular reflections on the settings. Simply removing specular reflections from the stimulus to isolate the influence of distortion information does not solve this issue because this in turn strongly reduces the impression of transparency (see Experiment 1b in this article). A second problem is how systematic deviations from a perfect match, as they were found in the experiments of Fleming et al. (2011b), should be interpreted. While such deviations could be attributed to the heuristic nature of visual perception or unfavorable stimulus conditions, they could as well indicate a fundamental flaw in the underlying theory. The interpretation of such deviations is especially difficult because the RI hypothesis – at least in its current form – is not specific enough to quantitatively predict the susceptibility of the estimates to variations in the input. It seems, therefore, necessary, to evaluate the RI hypothesis in a more indirect way that encompasses systematic tests of hypotheses derived from the theoretical assumptions, critical

comparisons of the experimental results with the predictions of alternative explanations, and theoretical analyses of computational and functional aspects. As a first step in this direction, we presented in Schlüter and Faul (2014) a theoretical analysis that leads to the conclusion that the RI hypothesis is implausible from a computational perspective. Furthermore, we found that the subjects' settings in the matching task were substantially influenced by irrelevant context factors.

In the present article, we test two predictions derived from the computational ideas that Fleming et al. (2011b) presented as a specification of their RI hypothesis (see Figure 3.2 for a schematic illustration). They first consider the *displacement field* D, a 2D vector field that "measures the displacement of all features in the background when seen through the transparent object" (p. 814): $D(x, y) = p_r - p_i$, with $p_i = (x, y)$ being the image location of a feature in undistorted plain view and $p_r = (x', y')$ being the location of the same feature when refracted by the transparent object. Second, they consider the divergence $d(x, y) = \nabla D(x, y)$ of the vector field D, which they call the *distortion field*. The divergence is a scalar field that represents only a part of the information contained in the displacement field D. Its values are related to the "relative magnitude of compression" (p. 814) of the texture pattern. The core assumption of Fleming et al. (2011b) is that it is somehow possible to compute an estimate $\hat{d}$ of the distortion field d from the input image and that this estimate could then be used in a second step to infer the refractive index. If both D and p_i are completely unknown, it is obviously impossible to estimate $\hat{d}$ from the information given in the distorted image p_r alone. This implies that in order for this computational idea to work, constraints for p_i are required. Fleming et al. (2011b) suggest that these constraints can be obtained by referring to the undistorted background, if it is visible in the surround of the object. More specifically, they propose that using the distortion field as a cue "involves comparing the relative scale of texture elements seen through the transparent object with the elements seen directly" (p. 818). In addition to the problems associated with estimating the distortion field, Fleming et al. (2011b) identify the further problem of "how the local estimates of distortion magnitude are pooled into a global estimate of RI" (p. 818). In our previous investigation of the RI hypothesis (Schlüter & Faul, 2014), we mainly focused on this RI estimation problem (i.e., step 2 in Figure 3.2) and argued that it would be computationally highly complex because of the many context factors that also influence the distortions but are not linked to the object's refractive index (e.g., the viewpoint, the object's shape or the distance of the background).

In this article, we concentrate on empirical tests of the first step, the estimation of the distortion field. To this end, we derive two testable predictions from the RI hypothesis that do not require any knowledge about the second step that presumably uses the estimate $\hat{d}$ to infer the object's refractive index. In essence, the first expectation is that the subjects' settings should be largely invariant against global changes in the background texture because this manipulation neither systematically influences the distortion field nor the relative scale

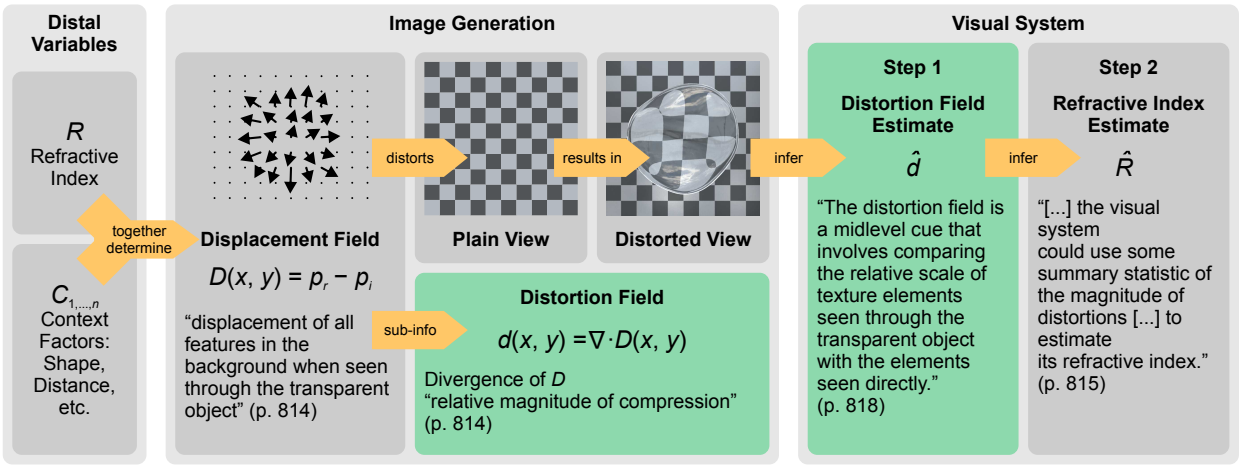

Figure 3.2: Schematic illustration of the RI hypothesis of Fleming, Jäkel, and Maloney (2011b), which states that the visual system can use image distortions caused by thick transparent objects to estimate the refractive index of the object's material. The authors describe the distortions by means of a 2D vector field, the displacement field D, that contains information about the direction and magnitude by which the position of each image feature is displaced in the distorted view (p_r) compared with the plain view (p_i). The displacement field depends on the refractive index of the material (R) and several context factors ($C_{1,...,n}$, e.g., object shape, background distance, viewpoint). Fleming, Jäkel, and Maloney (2011b) suggest that the visual system refers to the divergence d of the displacement field, which represents only part of the information available in D. They assume that this so-called distortion field d can be estimated by comparing the relative scale of texture elements within the object and the surround (Step 1). This distortion field estimate $\hat{d}$ is then supposed to be used to infer the refractive index estimate $\hat{R}$ (Step 2).

of the texture elements in the distorted and undistorted parts of the background. A second expectation is that the error in the subjects' settings should increase substantially if the putatively important information given in the object's surround is removed because the assumed computational strategy is then no longer possible.

We test both predictions and compare them with the predictions derived from the alternative explanation presented in Schlüter and Faul (2014), which assumes that the subjects do not match estimated refractive indices but instead simply maximize the similarity of the presented stimuli on the image level. This matching strategy is completely unrelated to transparency perception.

3.2 Experiment 1: The influence of background texture density

Experiment 1 tests the first prediction derived from the RI hypothesis that changes in scene variables that do not influence the displacement field should have no systematic effect on the refractive index settings. This test is described in Experiment 1a. In Experiment 1b, we compare the results of Experiment 1a with the predictions from our alternative explanation.

3.2.1 Experiment 1a: Testing the prediction of the RI hypothesis

The displacement field D and thus also the distortion field d depend on several variables, for example, the material and the shape of the object, its distance to the background, the viewing angle, but not on the background texture. In a scene that is geometrically fixed, D may be regarded as an operator that acts on arbitrary backgrounds and in each case yields the same distorted version of it. Thus, if the RI hypothesis is correct, then the refractive index estimate should be invariant against global changes in the background texture, as long as the texture contains enough structure to reflect the optical distortions truthfully.

To test this prediction, we performed an experiment similar to the ones conducted by Fleming et al. (2011b) and Schlüter and Faul (2014). The subjects saw two transparent objects of slightly different shape that were placed in front of backgrounds with different texture densities. All other scene variables were identical. The task was to match the material properties of a standard and a test object by adjusting the refractive index of the test (see Figure 3.3).

3.2.1.1 Stimuli

The stimulus material consisted of computer-generated images that were created with the Mitsuba renderer (Jakob, 2013). They were similar to the ones used in Schlüter and Faul (2014), which were designed to closely resemble the ones used by Fleming et al. (2011b). The

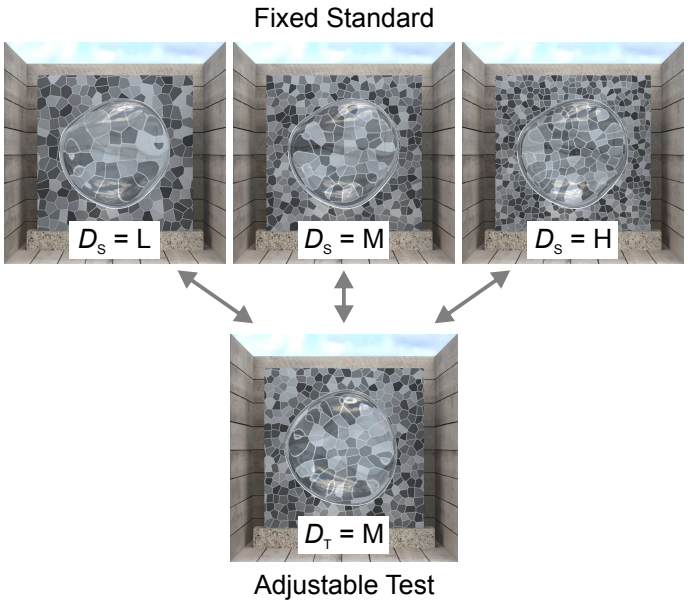

Figure 3.3: In the material matching task of Experiment 1a, a fixed standard stimulus, whose background texture density and refractive index varied according to the experimental condition, was compared with a test stimulus, whose refractive index was adjustable by the subjects. The texture density of the test stimuli (D_T) was medium ("M") throughout the experiment. The density of the standard stimuli (D_S) was identical ("M"), higher ("H"), or lower ("L") than in the test. The objects illustrated here have identical refractive indices ($R_S = R_T = 1.5$).

stimulus images showed a thick transparent object in front of a background plane that was located inside a box with front and top sides open (cf. Figure 3.3). All scene elements were defined in real-world coordinates relative to a virtual projection plane, which represented the surface of the experimental screen. Standard and test objects were two slightly different warped ellipsoids. The objects were modeled with the 3D computer graphics software Blender (Blender Foundation, 2013), by applying various shape modifiers to an icosahedron that was subdivided eight times. The objects were located at the center of the virtual projection plane, had a size of $50 \times 50 \times 6$ mm and were defined to be light-transmitting ("dielectric") without any absorption. The refractive index of the standard object varied in three steps ($R_S \in \{1.20, 1.50, 1.80\}$). The refractive index of the test object was adjusted by the subjects ($R_T \in \{1.010, 1.015, \ldots, 2.495, 2.500\}$, 299 steps). The refractive index of the surrounding medium was set to 1.

The textured background plane ($80 \times 80 \times 6$ mm) was placed behind the transparent object at a distance of 60 mm. The background textures were random Voronoi patterns created with "MATLAB and Statistics Toolbox" (2013) that resembled the background textures used by Fleming et al. (2011b). The individual faces of the pattern were separated by bright achromatic ($x = .30$, $y = .32$, $L = 52.75\,\mathrm{cd/m^2}$) seams of 0.32 mm width. The faces were also achromatic ($x = .31$, $y = .32$), and their luminances were uniformly distributed between 4.43 and

47.43 cd/m^2. The density of the background textures of the standard stimuli varied in three steps according to the experimental condition ($D_S \in$ {"L", "M", "H"}, see Figure 3.3). The medium density texture ("M") contained roughly 20×20 faces, the low density texture ("L") 15×15 faces and the high density texture ("H") 25×25 faces. The density of the background texture for the test stimuli was the same throughout the experiment (D_T = "M" , roughly 20×20 faces). It was computed with a different random seed than the texture of the same density used in the standard stimulus.

Image-based lighting with an infinitely distant spherical emitter was used as the sole illumination. The illumination texture was a high dynamic range image of a natural daylight outdoor scene with a partly cloudy sky. The camera settings (location and field of view) corresponded to the actual experimental setup. Thus, the stimuli appeared in nearly the same way as a corresponding real scene, except for the lack of binocular disparity. Stimuli were rendered as 16 bit/channel high dynamic range images ("extended volumetric path tracer"; "low discrepancy sampler" with 512 samples/px; Gaussian reconstruction filter with $SD = 0.5$) and subsequently tonemapped to 8 bit/channel low dynamic range images ($\gamma = 1.6$; exposure = 1.0). The final image size was 370×370 px which corresponded to 100×100 mm on the screen.

3.2.1.2 Procedure

In each trial, a fixed standard and an adjustable test stimulus were presented simultaneously on an Eizo ColorEdge CG243W LCD screen (display area 518.4×324.0 mm; resolution 1920×1200 px; color depth 8 bit/channel; 3.704 px/mm; Eizo Corporation, Hakusan, Japan), which was located in a darkened room. The standard stimulus was presented in the upper half of the screen, the test stimulus in the lower half. The viewing distance was 40 cm. The subjects' task was to adjust the test object until it appeared to be made of the same material as the fixed standard object. The refractive index of the test object was adjusted with the arrow keys of a standard computer keyboard. Each subject performed 27 trials in a pseudo-randomized order: $3R_S \times 3D_S \times 3$ repetitions.

This and all following experiments were carried out in accordance with the relevant institutional and national regulations and legislation and with the World Medical Association Helsinki Declaration as revised in October 2008.

3.2.1.3 Subjects

Seven subjects, six of them female, participated in the experiment. Their age ranged from 20 to 27. All subjects were naive as to the purpose of the experiment, except one, which was one of the authors (N.S.). They reported normal or corrected-to-normal visual acuity and showed no color vision deficiency, as tested by Ishihara plates Ishihara (1969).

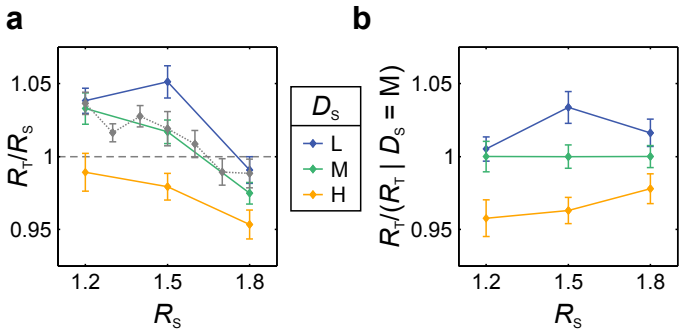

Figure 3.4: Results of Experiment 1a. **(a)** The adjusted refractive indices (displayed as test to standard ratios $R_T/R_S \pm SEM$) obtained with different background texture densities in standard and test (i.e., $D_S \in$ {"L", "H"}) deviate systematically from those obtained with an identical texture density (i.e., D_S = "M"). The settings for trials with the same texture density deviate only marginally from those found with uniform backgrounds (dotted gray line; reprinted from Schlüter and Faul, 2014, Figure 13). **(b)** The same test to standard ratios plotted relative to the ones obtained with identical texture density. This highlights the systematic deviations due to density differences. In both plots, each colored solid line shows the mean refractive index setting for a particular background texture density in the standard stimulus (D_S) as a function of the refractive index of the standard object (R_S).

3.2.1.4 Results

The refractive indices obtained with different background texture densities in standard and test deviate systematically from those obtained with identical texture densities (Figure 3.4a). The small deviations found with identical texture densities (i.e., D_S = "M") are very similar to those found previously with a uniform gray background under otherwise identical conditions (cf. Schlüter and Faul, 2014, Figure 13) that are given as a dotted gray line in the figure. In Figure 3.4b, the settings are plotted relative to the ones obtained with identical texture density. Relative to this condition, the adjusted refractive indices are larger for lower density in the standard (i.e., D_S = "L") and smaller for higher density (i.e., D_S = "H"). A two-way analysis of variance revealed a significant main effect of background texture density, $F(2, 180) = 22.33$, $p < .001$.

3.2.1.5 Discussion

The results obtained in Experiment 1a contradict the prediction derived from the RI hypothesis. Even moderately different background densities in otherwise similar stimuli lead to significantly different refractive index settings. It is less the absolute amount of these deviations but their systematic nature that is diagnostic because systematic deviations are inexplicable if the subjects' settings are based on the postulated distortion field. One cannot exclude the possibility that a manipulation of the texture density influences the accuracy of the distortion field estimation, but this should mainly affect the size of random errors. The random errors, however, are of comparable size in all conditions.

The settings obtained in the condition with identical texture densities in standard and test closely replicate previous results with completely uniform backgrounds (Schlüter & Faul, 2014). This suggests that the deviations from identity found in this condition are not related to the properties of the background texture but due to other factors, for instance, peculiarities in the specular reflection component resulting from the slightly different shapes of the standard and test objects.

Compared with the stimuli with a uniform background that were used in Schlüter and Faul (2014), the current stimuli contain background distortions which can, according to the RI hypothesis, also be used to estimate the refractive index. However, the addition of this cue did not only fail to improve the estimation accuracy but even impaired it. This is contrary to what is expected if the RI hypothesis holds.

3.2.2 Experiment 1b: Testing the prediction of the image-matching hypothesis

The alternative explanation presented in Schlüter and Faul (2014) assumes that the subjects do not compare estimated refractive indices in such experiments but simply maximize the similarity of the presented background textures on the image-level. Such image-level matches can potentially refer to any statistic of texture attributes. An important property of this kind of match is that it is completely unrelated to transparency perception. This distinguishes it from the refractive index match assumed by Fleming et al. (2011b).

It is at present unclear which information is used in this type of match. The settings that would result under the conditions of Experiment 1a must therefore be determined empirically. To this end, we isolated the background distortions in the stimuli used in Experiment 1a by removing all specular reflections (Figure 3.5, condition “Isol”). After this manipulation, the transparency impression is almost completely lost. For the present purposes, this is an advantage because this makes it highly plausible that the subjects actually perform image-level matches as intended.

A problem arising in this process is that areas of total reflection persist in the image as black areas at the object’s border. To prevent subjects from matching these black ring-like areas, whose width correlates with the refractive index, we masked the border area of the object with a gray ring (Figure 3.5, condition “IsolMsk”).

To be able to draw conclusions from the results of this experiment on the matching behavior in Experiment 1a, we kept most methodological aspects identical, including the instruction that asked the subjects to match the objects’ material properties. We suspected that this might confuse the subjects because the material impression in the manipulated stimuli is weak or even absent. To control for this, we restored a material impression by adding a fixed specular reflection pattern to the border of the objects (Figure 3.5, condition “IsolSR”), which

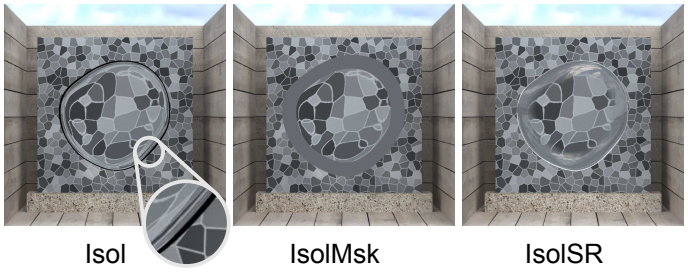

Figure 3.5: Stimulus conditions of Experiment 1b. We used the same stimuli as in Experiment 1a but isolated the influence of background distortions on the subjects' settings. In the condition "Isol", all specular reflections were omitted to isolate the information from background distortions. To hide rudiments of the specular reflections (black pixels at areas of former total reflections, see inset picture in the "Isol" stimulus), a ring shaped gray mask was added in the control condition "IsolMsk". To retain the impression of a material, and thus the plausibility of the material match instruction, we added static specular reflections to the border of the object in a second control condition ("IsolSR").

did not depend on the refractive index that was used to compute the background distortion in standard and test.

3.2.2.1 Stimuli

The stimuli were the same as those used in Experiment 1a, except for the following differences: For the condition "Isol", all specular reflections were omitted during the rendering process. The resulting darkening of the object area (according to Fresnel's equations, the amount of light that is transmitted through a surface decreases with increasing refractive index) was cancelled by correcting the average brightness of the object area to keep subjects from using it as a matching criterion. In the control condition "IsolMsk", remnants of the specular reflections (black ring-like areas that occurred at areas of former total reflections, cf. Figure 3.5) were masked by a gray ring ($x = .31$, $y = .31$, $L = 17.21\,\mathrm{cd/m^2}$). In the second control condition "IsolSR", fixed specular reflections were added to the inner object border by superimposing an image of an object with a constant refractive index ($R = 1.5$) in front of a uniform background texture ($x = .31$, $y = .32$, $L = 17.21\,\mathrm{cd/m^2}$). The inward transition to the original stimulus was blurred.

3.2.2.2 Procedure

The procedure was the same as in Experiment 1a. However, each subject performed 81 trials in a pseudo-randomized order ($3R_S \times 3D_S \times 3$ isolation conditions $\times$ 3 repetitions).

3.2.2.3 Subjects

The subjects were the same as in Experiment 1a.

3.2.2.4 Results

For all conditions, the refractive index settings deviate systematically from those given in the standard object (Figure 3.6a). If the background texture density of the standard is lower than that of the test (i.e., D_S = "L"), the refractive index settings are higher than the given ones ($R_T/R_S > 1$), while they are lower for a higher density in the standard (i.e., D_S = "H"). The deviations are comparatively small if standard and test have the same texture density (i.e., $D_S = D_T$ = "M"). The size of the deviations in the two control conditions ("IsolMsk" and "IsolSR") is similar and slightly larger than in the "Isol" condition.

These differences become especially apparent if the settings are averaged across all R_S (Figure 3.7, left three columns). If the settings are directly compared with the ones gained in Experiment 1a (Figure 3.7, fourth column), it becomes apparent that the general pattern of deviations is the same in both experiments, but that the size of the deviations is much larger in Experiment 1b. The settings for stimuli with uniform backgrounds that were reported in Schlüter and Faul (2014) are depicted for comparison (see discussion).

3.2.2.5 Discussion

In this experiment, the subjects had only the (distorted) backgrounds available for their matches. Figure 3.8 shows the subjects' mean test settings for standards with different texture densities. Comparing the background patterns suggests that the subjects matched the average element size of the distorted background across the scenes. A low background texture density in the standard results in a large average element size within the distorted background. To match this, the test's refractive index has to be increased above the standard's one until the magnification effect leads to a similar average element size (roughly speaking, increasing refraction increases the magnifying effect of the convex light-transmitting object, which in turn increases the average element size). Mutatis mutandis, a standard with a higher background texture density leads to the prediction of a deviation in the opposite direction.

The similarity of the pattern of deviations found in Experiments 1a and 1b suggests that image-level matches of the distorted background textures had also played a critical role in Experiment 1a. The considerably smaller size of the deviations in Experiment 1a can be explained by assuming that the subjects also used information from specular reflections. Isolated matches of specular reflections have been found to lead to settings that are almost identical to the ones given in the standard (see replot of data from Schlüter and Faul, 2014, in Figures 3.4 and 3.7). Because both types of matches require different adjustments of the refractive index, the subjects likely settled on a compromise. This would explain the shift toward identity with increasing relative salience of reflection-related information.

This interpretation is also supported by the finding that the deviations are lower in the "Isol" condition than in the two control conditions because the traces of information from

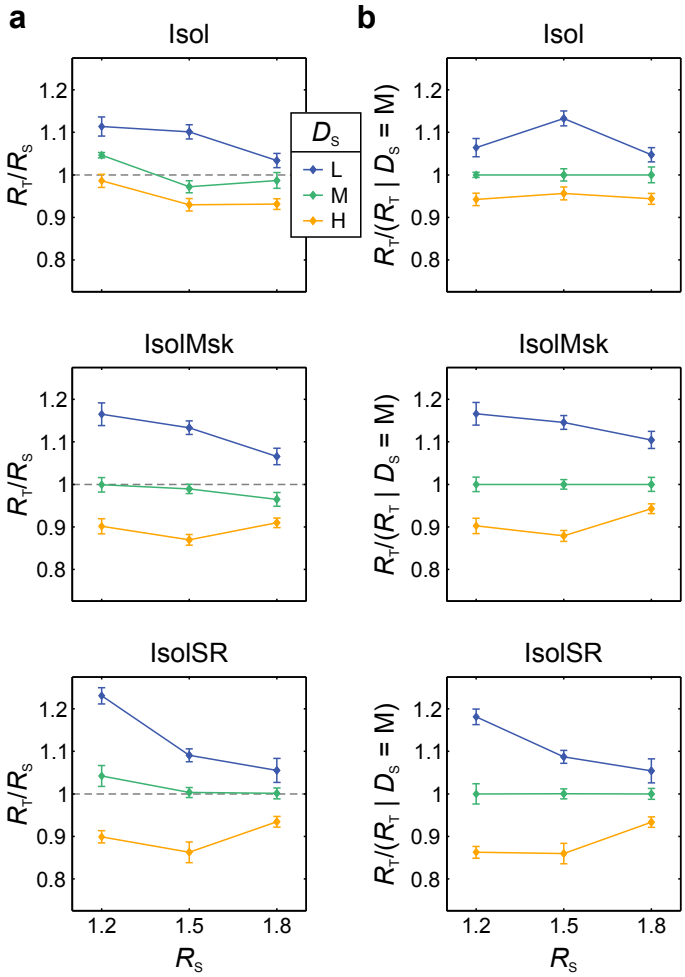

Figure 3.6: Results of Experiment 1b. **(a)** The refractive index settings (displayed as test to standard ratios $R_T/R_S \pm SEM$) obtained with different background texture densities in standard and test (i.e., $D_S \in$ "L", "H") deviate systematically from those obtained with identical texture density in standard and test (i.e., $D_S =$ "M"). The pattern of deviations is similar in all conditions, but their size is slightly smaller in the "Isol" condition than in the two control conditions "IsolMsk" and "IsolSR". **(b)** To highlight the systematic deviations that can be attributed to density differences, the same test to standard ratios are plotted relative to the ones obtained with identical density in standard and test (i.e., $D_S =$ "M"). Each colored solid line shows the mean refractive index setting for a particular background texture density in the standard stimulus (D_S) as a function of the refractive index of the standard object (R_S).

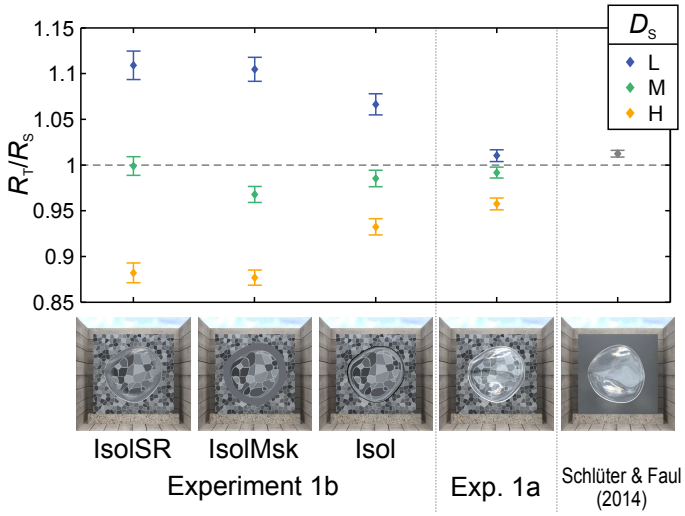

Figure 3.7: The results of Experiments 1b and 1a in comparison with the settings for a uniform background reported in Schlüter and Faul (2014). While the general pattern of deviations is similar in Experiments 1a and 1b, the deviations differ in size and are considerably larger in Experiment 1b. The settings obtained in matches of isolated specular reflections are close to identity. These results suggest that the size of the deviations decreases as the relative salience of reflection-related information increases (from "IsolSR" on the left to the uniform background on the right). Points of the same color correspond to the mean test to standard ratio ($R_T/R_S \pm SEM$) for a particular background texture density in the standard stimulus (D_S) averaged across all refractive indices of the standard object (R_S).

specular reflections available in the "Isol" condition may also have shifted the subjects' settings closer towards identity. The smaller degree of this shift might be explained by the plausible assumption that the "black ring" cue is less salient than specular reflections and therefore is given a lower weight in the compromise.

The similarity of the results in the "IsolSR" and the "IsolMsk" condition suggests that the subjects were not confused by the fact that we retained the material matching instruction although the impression of an object with distinct material properties is rather weak or even absent in the "Isol" and "IsolMsk" conditions.

3.3 Experiment 2: The role of the surround

Experiment 2 is related to the assumption of the RI hypothesis that estimating the distortion field requires a comparison of distorted and undistorted background areas. In Experiment 2a, we tested a prediction derived from this assumption. In Experiment 2b and 2c we compare these results with predictions from our alternative explanation.

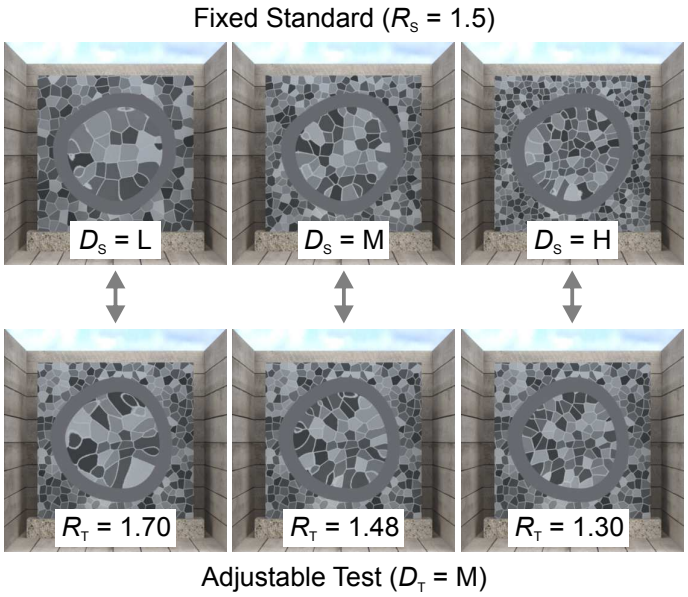

Figure 3.8: Exemplary depiction of the subjects' mean settings in the test object (bottom row) in the control condition "IsolMsk" for standard stimuli (top row) with a fixed refractive index ($R_S = 1.5$) and a background texture density (D_S) that was either lower (left column, "L"), the same (middle column, "M"), or higher (right column, "H") than the one of the test stimulus (D_T = "M"). If standard and test had different background texture densities, the adjusted refractive indices (R_T) differed systematically from the fixed ones given in the standard (R_S). The depicted results suggest that the subjects tried to equate the average element size of the distorted background textures.

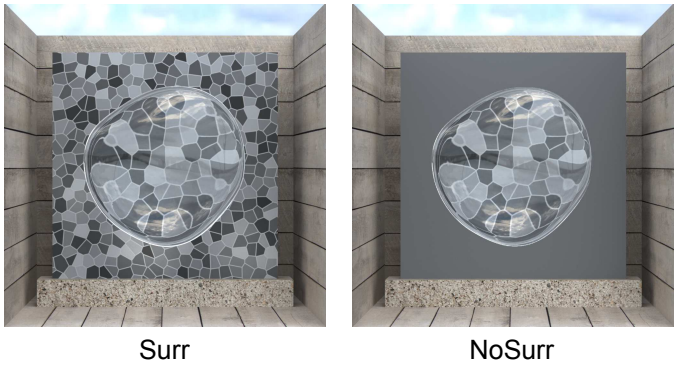

Figure 3.9: Stimulus conditions of Experiment 2a. Because estimating the distortion field presumably requires a comparison of the distorted background within the object boundaries with the undistorted background in the surround, the subjects' settings should become highly unreliable if the surround is removed. To test this prediction, we used similar stimuli as in Experiment 1a. The surround was either left unchanged ("Surr"), or it was replaced by a uniform gray background ("NoSurr").

3.3.1 Experiment 2a: Testing the prediction of the RI hypothesis

Fleming et al. (2011b) assume that using the distortion field as a cue "involves comparing the relative scale of texture elements seen through the transparent object with the elements seen directly" (p. 818). This implies that the estimation of the refractive index from the distortion field should be severely hampered or may even fail completely if the texture in the object's surround is removed. As a consequence of this manipulation, one would expect highly unreliable settings and accordingly a strong increase in the size of (random) errors. Alternatively, it would also be possible that the putative distortion field cue is identified as unreliable and simply ignored. In that case, one would expect settings similar to those obtained in Schlüter and Faul (2014) for stimuli with a uniform background, that is, approximate identity of the refractive indices in standard and test. To test these predictions, we repeated Experiment 1a with an additional condition in which the texture in the surround was replaced with a uniform gray area (condition "NoSurr", Figure 3.9).

The subjects' matching behavior may depend on the order in which the stimuli with and without a textured surround are presented to them. For example, the subjects might switch to an alternative matching strategy when the textured surround disappears for the first time and stay with it for the sake of consistency even if the textured surround becomes available again in a following trial. To control for such effects, we balanced the order with which the two surround conditions were presented to the subjects. Half of the subjects had the "Surr" condition first, the other half the "NoSurr" condition ("Surr →NoSurr") vs. "NoSurr →Surr").

3.3.1.1 Stimuli

The stimuli were similar to those used in Experiment 1a. Again, the background texture density of the standard stimuli was varied in three steps ($D_S \in$ "L", "M", "H"), while the test stimuli always had the same density (D_T = "M"; cf. Figure 3.3). The undistorted background visible in the surround of the object was either left unchanged ("Surr") or was replaced by a uniform gray background ("NoSurr"; $x = .31$, $y = .32$, $L = 17.21\,\mathrm{cd/m^2}$) as depicted in Figure 3.9.

3.3.1.2 Procedure

The procedure was the same as in Experiment 1a. However, each subject performed 54 trials ($3R_S \times 3D_S \times 2$ surround conditions $\times$ 3 repetitions). In each half (i.e., "Surr", "NoSurr") of the conditions "Surr →NoSurr" and "NoSurr →Surr", the trials were presented in random order.

3.3.1.3 Subjects

Eight subjects, six of them female, participated in the experiment. Their age ranged from 19 to 37. All subjects were naïve as to the purpose of the experiment and had not participated in similar experiments. They reported normal or corrected-to-normal visual acuity and showed no color vision deficiency, as tested by Ishihara plates (Ishihara, 1969).

3.3.1.4 Results

The subjects' settings do not depend systematically on the availability of the textured surround ("Surr" vs. "NoSurr", Figure 3.10) or the order by which the surround conditions were presented ("Surr →NoSurr" vs. "NoSurr →Surr", Figure 3.11). The settings of the refractive indices appear almost identical to those observed in Experiment 1a. For standard stimuli with lower background texture densities (D_S = "L"), the refractive index settings are higher than the given ones ($R_T/R_S > 1$), for those with higher background texture densities (D_S = "H"), the settings are lower ($R_T/R_S < 1$). The deviations are considerably smaller if standard and test stimuli have the same background texture density (i.e., $D_S = D_T$ = "M").

3.3.1.5 Discussion

The prediction derived from the RI hypothesis that the errors of the refractive index settings will increase if the undistorted background texture in the surround is removed is not confirmed by our data. Instead, both the pattern of systematic deviations and the size of random errors are essentially identical for conditions with and without a textured surround. This suggests that the subjects referred solely to information within the object boundaries and ignored any information available in the surround.

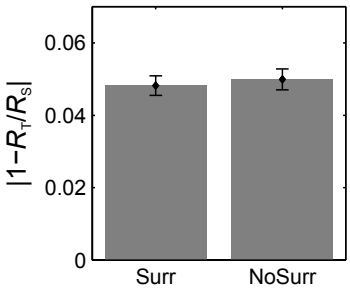

Figure 3.10: Mean absolute deviations of the adjusted refractive indices (displayed as test to standard ratios $R_T/R_S \pm SEM$) from identity (i.e., $R_T/R_S = 1$) for stimuli with ("Surr") and without ("NoSurr") textured surround.

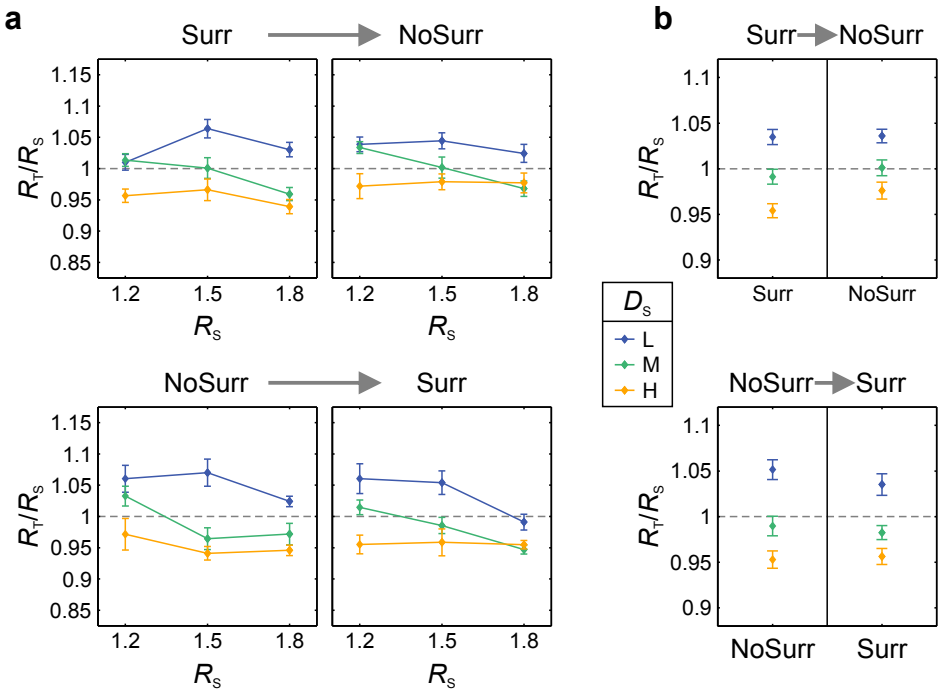

Figure 3.11: Results of Experiment 2a. Neither the availability of the surround nor the presentation order has a systematic effect on the subjects' settings. **(a)** Mean refractive index setting (displayed as test to standard ratios $R_T/R_S \pm SEM$) for a particular background texture density in the standard stimulus (D_S) as a function of the refractive index of the standard object (R_S). Stimuli without a textured surround ("NoSurr") were either shown after ("Surr →NoSurr", top row) or before ("NoSurr →Surr", bottom row) stimuli with a textured surround ("Surr"). **(b)** The same data averaged across all R_S.

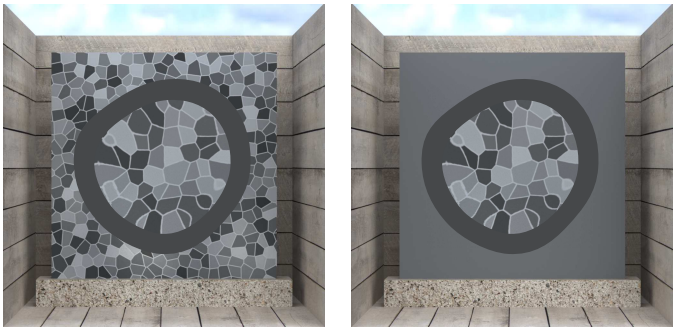

Figure 3.12: Stimulus conditions of Experiment 2b. We used the same stimuli as in Experiment 2a but isolated the influence of background distortions on the subjects' settings. If subjects only refer to the object's area, then the availability of the textured surround should not have an influence on their settings.

Furthermore, our results show that the subjects did not ignore the remaining background texture completely, since their settings differ from those obtained with a uniform background (Schlüter & Faul, 2014).

These observations do not seem to be compatible with the computational idea underlying the RI hypothesis.

3.3.2 Experiment 2b: Testing the prediction of the image-matching hypothesis

In Experiment 2b, we again isolated the background distortions to determine the settings resulting from image-level matches under the conditions realized in Experiment 2a. If the assumption is correct that the observed deviations in Experiments 1a, 1b, and 2a are the result of image-level matches of properties of the distorted textures, then the following results would be expected: (a) The pattern of the deviations observed in Experiments 2a and 2b should be identical because the availability of the textured surround is then irrelevant for the settings. (b) The absolute size of the deviations should be larger in Experiment 2b than in 2a because the compensating effect of the reflection cue is no longer available (see discussion of Experiment 1b). (c) The results of Experiments 1b and 2b should be similar because under the above assumption the two conditions were equivalent.

We only used a single isolation condition ("IsolMsk", see Experiment 1b), and the surround was manipulated in the same way as in Experiment 2a (see Figure 3.12).

3.3.2.1 Stimuli

The stimuli were identical to those used in Experiment 2a, except that background distortions were isolated in the same way as in the "IsolMsk" condition of Experiment 1b, that is, any specular reflections were omitted during the rendering, the brightness of the object area was corrected for any refraction-dependent darkening, and a gray mask was added to the object boundary ($x = .31$, $y = .32$, $L = 7.88\,\mathrm{cd/m^2}$).

3.3.2.2 Procedure

The procedure was the same as in Experiment 2a. Again, each subject performed 54 trials ($3R_S \times 3D_S \times 2$ surround conditions $\times$ 3 repetitions).

3.3.2.3 Subjects

The subjects were the same as in Experiment 2a.

3.3.2.4 Results

The results of Experiment 2b are shown in Figures 3.13a and 3.13b. Overall, the settings in trials with different texture densities in standard and test deviated from identity in a similar way as in Experiment 1b, as expected. However, a completely different pattern of results is observed in the second half of the "NoSurr⟶Surr" condition: For three out of four subjects the deviations almost vanish and the refractive index settings virtually coincide with each other, irrespective of the actual background texture density. For one of the subjects, this collapse does not occur. The results of this subject are displayed separately in Figure 3.13b (labelled "$n = 1$").

3.3.2.5 Discussion

In general, the pattern of deviations in this experiment is very similar to the one found in Experiment 2a. This suggests that image-level matches of background distortions had also played a critical role in Experiment 2a.

An unexpected observation is the "collapse of the deviations" in the second half of the "NoSurr ⟶Surr" condition. These settings are close to identity and one may therefore be tempted to interpret this as support for the RI hypothesis. However, several observations speak against this view: First, large systematic deviations from identity were found if the exact same stimuli were presented at the beginning of the experiment in the "Surr ⟶NoSurr" condition. Second, the collapse of deviations was found in a physically inconsistent situation, in which the "objects" did not even appear transparent, whereas in Experiment 2a, in which the stimuli

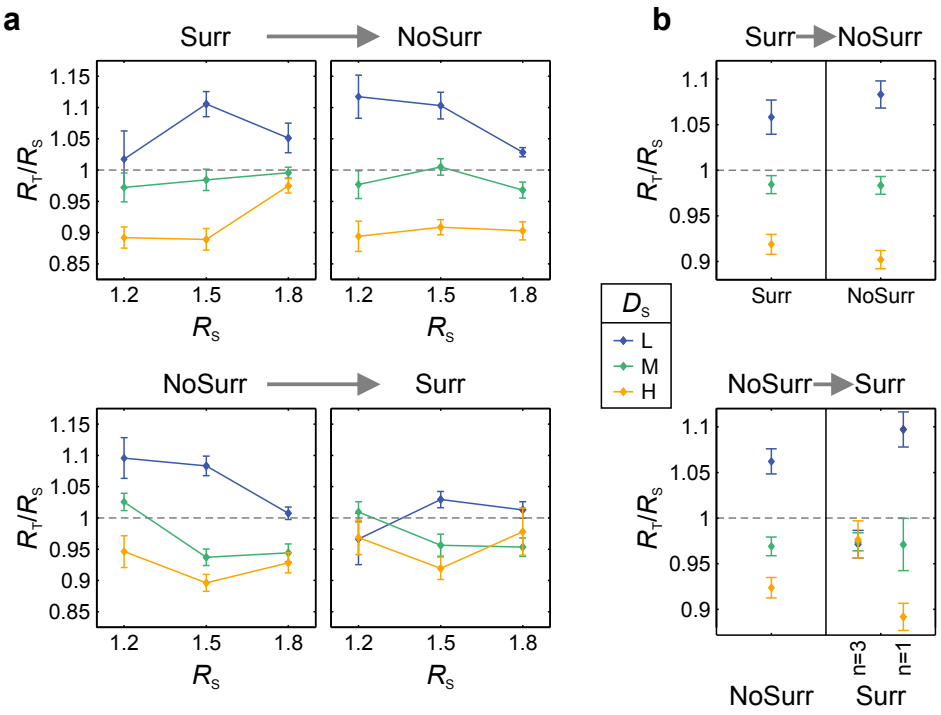

Figure 3.13: Results of Experiment 2b. The overall pattern of deviations is similar to the one found in Experiment 2a. However, if the textured surround is first shown in the second half of the experiment ("NoSurr →Surr"), then the settings almost coincide with each other, irrespective of the standard's background texture density (DS). **(a)** Mean refractive index setting (displayed as test to standard ratios $R_T/R_S \pm SEM$) for a particular background texture density in the standard stimulus (D_S) as a function of the refractive index of the standard object (R_S). In the condition "Surr →NoSurr" (top row), the textured surround was shown in the first half of the experiment, in the condition "NoSurr →Surr" (bottom row) in the second half. **(b)** The same data averaged across all RS. The collapse of deviations for "Surr" stimuli shown in the second half of the experiment ("NoSurr →Surr") occurs for three out of four subjects (labelled "$n = 3$"). The settings of the remaining subject are given separately (labelled "$n = 1$"). Please note that the panel on the left corresponding to this condition shows the mean across all four subjects.

contained at least two consistent and physically correct cues and appeared transparent, no such collapse was observed. Third, not all subjects showed this collapse of deviations.

A more plausible explanation is that the subjects in this case matched relations between image attributes, for instance, the ratio of the mean element size within and outside of the object's area, instead of image attributes within the object's boundary directly, as in the other conditions. In the particular situation realized in the experiment, this can lead to settings near identity because these relations are presumably almost invariant against changes in the background texture density. The observed result is to be expected if three of the four subjects had deliberately switched their matching strategy from absolute to relational, when the surround suddenly changed from uniform to textured in the second half of the experiment. It is at present unclear, why such a switch did not occur in Experiment 2a, but it seems plausible to assume that the salient specular reflections that also influence the settings kept the subjects focused on the object areas.

3.3.3 Experiment 2c: Testing relational image-level matches

To test, whether the collapse of deviations in Experiment 2b could be explained by a match of relations between image-attributes, we explicitly instructed the subjects in a similar matching experiment to perform this kind of match.

We also tried to address the question, whether such relational matches could constitute a viable strategy (in the sense of Fleming et al., 2011b) to reliably estimate the refractive index of transparent objects. For this to be the case, the estimates need to be robust against changes in context. To test this, we varied the thickness of the standard object, that is, a context factor unrelated to material properties that also influences background distortions.

3.3.3.1 Stimuli

The stimuli were similar to the ones used in Experiment 2b. However, the thickness of the standard object was varied in three steps ($T_S \in \{3, 6, 9\,\text{mm}\}$), while the thickness of the test object remained constant ($T_T = 6\,\text{mm}$). Different thicknesses were implemented by applying a directional scaling factor to the meshes of the objects. In contrast to Experiment 2b, only stimuli with a textured surround ("Surr") were used, for which the "collapse of deviations" were found.

3.3.3.2 Procedure

The procedure was the same as in Experiment 2b. However, each subject performed 81 trials in randomized order ($3R_\text{S} \times 3D_\text{S} \times 3T_S \times 3$ repetitions). The subjects were explicitly instructed to refer with their matches to the relations of distorted and undistorted background areas

("Your task is to adjust the test stimulus until the size ratio of the texture elements within the object area and its surround is the same as in the standard stimulus.").

3.3.3.3 Subjects

Eight subjects, six of them female, participated in the experiment. Their age ranged from 21 to 29. All subjects were naïve as to the purpose of the experiment and did not participate in the previous experiments. They reported normal or corrected-to-normal visual acuity, and showed no color vision deficiency, as tested by Ishihara plates (Ishihara, 1969).

3.3.3.4 Results

The results of Experiment 2c are shown in Figures 3.14a and 3.14b. For all object thickness combinations, the settings do not depend on the background texture density. However, the settings are only near identity if standard and test objects have the same thickness. If the objects' thickness differs, the settings deviate systematically.

3.3.3.5 Discussion

The settings for objects of equal thicknesses resemble the "collapse of deviations" found in Experiment 2b. This suggests that the subjects' matching behavior was identical, despite the difference in the instructions. A plausible and parsimonious explanation of this finding is that the subjects in both experiments matched relations of texture properties between the distorted and undistorted background. An inspection of the match results suggests that the subjects referred to simple image attributes like the average element size of the background pattern. The ratio of such attributes taken between object area and surround is approximately invariant against changes in the overall background texture density. This explains the insensitivity of relational matches to differences in background texture density. Changes in the thickness of the object, in contrast, influence only the attributes within the object's boundaries (roughly speaking, increasing the thickness of the convex object increases its magnifying effect, which in turn increases the average element size in the background area seen through the object). This explains the systematic deviations found with objects of different thicknesses. These latter results are in conflict with the assumption that referring to the relation of simple image criteria is generally sufficient to estimate the distortion field and the refractive index.

3.4 Summary and general discussion

In this article, we tested two predictions derived from the RI hypothesis of Fleming et al. (2011b), which essentially states that the visual system can use image distortions caused by thick transparent objects to estimate the refractive index of their material. Fleming et

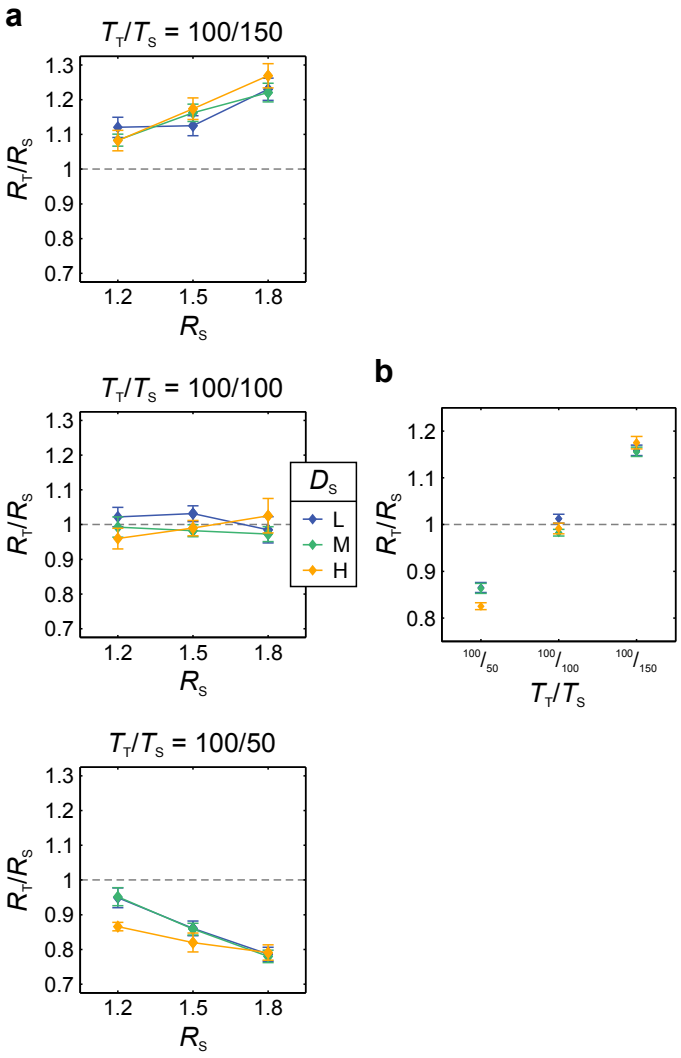

Figure 3.14: Results of Experiment 2c in which the subjects were instructed to match the relations of distorted to undistorted background areas. The refractive index settings (displayed as test to standard ratios $R_T/R_S \pm SEM$) depend on the object thickness (T_S) but not on the texture density (D_S) in the standard. **(a)** Each plot corresponds to one thickness of the standard object. Data points of same color show the mean refractive index settings for a particular background texture density in the standard stimulus (D_S) as a function of the refractive index of the standard object (R_S). **(b)** The same data averaged across all R_S.

al. (2011b) propose that this estimation process comprises two stages. In the first stage, a distortion field is estimated from the input pattern, and in the second stage, the information contained in this mid-level cue is "pooled into a global estimate of refractive index" (p. 818). The exact processes were largely left unspecified by the authors. However, it is thought that the estimated distortion field approximates some important aspects of the divergence of the displacement field, which describes the actual optical distortion. It is obvious that any kind of approximation would be impossible without some information about the undistorted background. Fleming et al. (2011b) therefore assume that the estimation of the distortion field "involves comparing the relative scale of texture elements seen through the transparent object with the elements seen directly" (p. 818).

Our strategy to test this computational idea was to derive predictions from it that could be tested without any knowledge about the exact nature of the putative processes. A first prediction is that the refractive index settings in a material matching task should not be influenced by isolated changes of a scene variable that is unrelated to the refractive index and does neither alter the relative scale of the texture elements inside the object and the surround nor the distortion field. In Experiment 1a, we tested this by asking subjects to match the material of two transparent objects across scenes that were identical apart from the density of their background texture. We found that even relatively small density differences had a highly systematic and statistically significant effect on the subjects' settings. Fleming et al. (2011b), in contrast, varied context factors (i.e., background distance and object thickness) that systematically influenced both the distortion field and the relative scale of texture elements within the object and its surround. The observed systematic deviations from a perfect match may thus be explained by assuming that the visual system cannot completely distinguish between different influences on the distortion field. Our strategy to keep the distortion field and the relative scale of texture elements constant excludes this kind of explanation and therefore suggests that the deviations observed in Experiment 1a are not due to a bias in the estimation of the refractive index, but that the matches were performed without referring to the distortion field at all.

The idea that the estimation of the distortion field involves a comparison of texture elements seen through the transparent object with elements seen directly leads to the further expectation that the uncertainty of refractive index settings that are based on this cue should increase sharply, if information about the undistorted background is removed. Contrary to this prediction, we found in Experiment 2a that the refractive index settings did not change when the texture in the surround was replaced by a uniform gray surface. Instead, we observed the same systematic deviations as in Experiment 1a. This suggests that the surround was completely ignored and that a comparison of distorted and undistorted background areas, as it has been proposed by Fleming et al. (2011b), did not take place. In general, it seems impossible to estimate the distortion field without using the surround as a reference for the undistorted

state of the background texture. While some information about the undistorted state may be inferred if rather strong regularity assumptions with respect to possible background patterns are made, for example, that they only contain straight lines, it would still be impossible to determine the absolute amount of the distortion. Such relative distortion information is insufficient to estimate the absolute value of the refractive index that uniquely characterizes a light-transmitting material. However, relative distortion information might be useful for other purposes, for example, to infer shape properties. Here, regions of relative magnification and compressions might serve as cues for convex or concave surface curvatures.

The results of Experiments 1a and 2a are clearly not in line with basic assumptions underlying the general computational idea outlined earlier. In three additional experiments, we tested whether the pattern of results found in Experiments 1a and 2a is compatible with simple image-level matches of the distorted background texture. The procedures and stimuli used in Experiment 1b and 2b were essentially identical to those in Experiment 1a and 2a, respectively. The sole difference was that we removed specular reflections and in this way isolated the information contained in the background distortions. This manipulation had the side effect that the "objects" did no longer appear transparent, which is an advantage in that it increases the plausibility that the subjects actually performed image-level matches. In these experiments, we observed deviations from identity that are larger in absolute value but exhibit the same pattern as in the corresponding experiments with physically plausible transparent objects. This is compatible with the interpretation that such image-level matches were also made in Experiments 1a and 2a, but that the presence of specular reflections shifted the settings closer towards identity. In Experiment 2b, we unexpectedly found "perfect matches" in a condition, where a textured surround first appeared in the second half of the experiment. The conjecture that this resulted from image-level matches of the relation of texture elements within and outside the objects' boundary was confirmed in Experiment 2c. The results obtained in this experiment further indicate that such relational image-level matches are in general not a viable strategy to estimate refractive indices.

3.5 Conclusions

The results obtained in Experiments 1a and 2a seem not compatible with fundamental principles postulated in the RI hypothesis and thus confirm the results and conclusions reported in Schlüter and Faul (2014). All results obtained so far in putative material matches of thick transparent objects can be more parsimoniously explained by the assumption that the subjects maximized the perceptual similarity of the presented stimuli on the image-level. They appear to use image criteria that vary with the adjustable refractive index parameter (e.g., the average element size of the background areas, the brightness of specular reflections, or the width of black artefacts). In the context of transparency perception, we consider such image-level

matches as a methodological artefact that has no theoretical meaning. Therefore, we have not examined in detail to which image criteria the subjects actually referred. (This question might however be of interest in the field of texture perception, e.g., Balas, 2006; Beck, 1983; Bergen & Adelson, 1988; Landy & Graham, 2004; Rosenholtz, 2015.) If our interpretation is correct, then it is to be expected that the subjects will refer to other image criteria in other kinds of stimuli and that they will be unable to perform reliable matches, if they fail to find an image property that systematically varies with the adjusted parameter.

Our findings seem also instructive from a methodological point of view because they highlight potential pitfalls in "direct" tests of hypotheses. The available evidence suggests that subjects in the material matching task of Fleming et al. (2011b) are unable to do what they are asked to do because the visual system cannot estimate the refractive index from image distortions in that situation. The important point here is that they do not refuse to perform a match, but instead tacitly switch to an unintended alternative task. Unfortunately, such a tacit change in task is notoriously difficult to detect because the resulting data are often highly reliable and may even partly be in line with the experimenter's expectations. The experiment of Fleming et al. (2011b) is a point in case because simple image-level matches of background distortions or specular reflections can lead to high correlations between the refractive index settings and the refractive index given in the standard and this can easily be mistaken as the result of refractive index matches. The mere fact that subjects were presented with transparent objects, adjusted refractive indices and were asked to match the objects' material does not guarantee that conclusions about transparency perception can be drawn from the obtained results.

3.6 Outlook

The theoretical considerations and empirical findings presented in Schlüter and Faul (2014) and the present article strongly suggest that observers are unable to estimate the refractive index from optical distortions in experimental settings similar to the ones realized by Fleming et al. (2011b). Strictly speaking, this conclusion is limited to the situations realized in these investigations. One may argue that the stimuli used in these experiments did not contain enough information to estimate and match the objects' refractive indices and that this forced the subjects to perform image-level matches. This hypothesis can – and should – be tested with more complex stimuli that provide additional information, for example, stereo and motion cues that may support the estimation of the distortion field and that facilitate the adequate consideration of irrelevant context factors. Unfortunately, such investigation will also be confronted with the difficult problem of how to control for the effects of specular reflections and other potential information related to the refractive index that may be available in more complex scenes. Simply removing additional information seems problematic because

this may reduce or even destroy the transparency impression. This was unproblematic in our Experiments 1b and 2b but is clearly undesirable if one tries to find support for the RI hypothesis. Even if it is possible to discard secondary cues without destroying the impression of transparency, this could nevertheless change the role the (isolated) cue of interest plays, depending on how the visual system actually integrates different cues (cf. Ernst & Bülthoff, 2004; Landy, Maloney, Johnston, & Young, 1995).

Whether the RI hypothesis can be confirmed in more complex scenes is an open empirical question. However, on theoretical grounds, we are not very optimistic about a positive outcome. In Schlüter and Faul (2014), we mentioned several reasons why we think that the RI hypothesis is a priori implausible. Apart from serious computational difficulties inherent in estimating refractive indices from background distortions, it remains unclear of what use such estimates could be, especially because it is seems probable that they cannot be very reliable.

A serious problem for tests of the RI hypothesis in its present form is that the specification of putative mechanisms is rather vague. It would be desirable to develop proposals that are more specific and to investigate with the help of computer simulations, whether and, if so, under which conditions they are sufficient to solve the problem of estimating the refractive index.

Kapitel 4

Visual shape perception in the case of transparent objects

Eine gekürzte Fassung dieses Kapitels wurde veröffentlicht als: Schlüter, N. & Faul, F. (2019). Visual shape perception in the case of transparent objects. *Journal of Vision*, *19*(4), 1–36. https://doi.org/10.1167/19.4.24

Zusammenfassung

Um die Form von Objekten zu schätzen, muss sich das visuelle System auf Regularitäten im (retinalen) Bild stützen, die mit der Form zusammenhängen. Für opake Objekte wurden bereits viele solcher Regularitäten identifiziert, die sich allerdings nicht ohne Weiteres auf transparente Objekte übertragen lassen, da sie dort entweder gar nicht oder nur in abgewandelter Form vorliegen. Wir betrachten hier drei potentielle Hinweise auf die Form, die spezifisch für transparente Objekte sind: optische Hintergrundverzerrungen durch Lichtbrechung, Farb- und Helligkeitsveränderungen durch Absorption und mehrfache Spiegelbilder der Umgebung durch spiegelnde Reflexionen. In Computersimulationen analysieren wir zunächst, in welchen Situationen diese Regularitäten als Formhinweis verwendet werden könnten. Anschließend untersuchen wir experimentell, wie die Formwahrnehmung vom Vorhandensein dieser potentiellen Formhinweise in realistischen Szenen unter natürlichen Beobachtungsbedingungen abhängt. Unsere Ergebnisse zeigen, dass die Form von transparenten Objekten sowohl weniger korrekt als auch weniger präzise wahrgenommen wurde als im opaken Fall. Darüber hinaus variierten die Stärke und Richtung des Einflusses einzelner Bildregelmäßigkeiten in Abhängigkeit der Objekt- und Szeneneigenschaften erheblich. Unsere Ergebnisse deuten darauf hin, dass es im transparenten Fall von einem komplexen Zusammenspiel der Eigenschaften des transparenten Objekts und der umgebenden Szene abhängt, welche Informationen als Formhinweis verwendet werden können.

Abstract

In order to estimate the shape of objects, the visual system must refer to shape-related regularities in the (retinal) image. For opaque objects, many such regularities have already been identified, but most of them cannot simply be transferred to transparent objects, because they are not available there at all or only in a substantially modified form. We here consider three potentially relevant regularities specific to transparent objects, namely optical background distortions due to refraction, changes in chromaticity and brightness due to absorption, and multiple mirror images due to specular reflection. Using computer simulations, we first analyze under which conditions these regularities may be used as shape cues. We further investigate experimentally how shape perception depends on the availability of these potential cues in realistic scenes under natural viewing conditions. Our results show that the shape of transparent objects was perceived both less accurately and less precisely than in the opaque case. Furthermore, the influence of individual image regularities varied considerably depending on the properties of both object and scene. This suggests that in the transparent case it depends on a complex interplay of properties of the transparent object and the surrounding scene what kind of information is usable as a shape cue.

4.1 Introduction

Perceiving the spatial extent and shape of objects is a fundamental ability that allows us to identify objects and to successfully interact with them. How the visual system infers the shape of opaque objects has already been investigated in a large number of theoretical and empirical works and several essential shape cues were identified, for example contours and edges, texture, shading, mirror images, and highlights (see Figure 4.1a). However, there are only a few studies that investigated visual shape perception in the case of transparent objects (e.g. Chen & Allison, 2013; Chowdhury, Marlow, & Kim, 2017; Interrante, Fuchs, & Pizer, 1995; Interrante, Fuchs, & Pizer, 1997; Kersten, Stewart, Troje, & Ellis, 2006; Wijntjes, Vota, & Pont, 2015). Most work on perceptual transparency deals with the transmission properties of simple flat filters and the color or brightness relations in the image that lead to perceptual transparency (Anderson, 2015; Beck, 1978; Beck, Prazdny, & Ivry, 1984; Faul, 2017; Faul & Ekroll, 2002, 2011, 2012; Faul & Falkenberg, 2015; Khang & Zaidi, 2002a, 2002b; Ripamonti, Westland, & Da Pos, 2004; Robilotto, Khang, & Zaidi, 2002).

The question arises whether the shape cues identified in the opaque case can somehow be transferred to the case of transparent objects. If one considers the substantial differences in light transport between those two material classes, this appears at least questionable. Light interacts with transparent objects in a much more complex way than with opaque objects: It can interact several times with the surface of the object, its interior, and the environment

before it reaches an observer. Depending on the actual material properties, this can not only change the spectral properties of the light but also its direction of propagation. These differences in light transport in the opaque and in the transparent case lead to differences in the information available in the image that may potentially be used as shape cues. Thus, the shape cues identified in the opaque case cannot simply be transferred to transparent objects.

In this work, we perform both theoretical analyses and empirical investigations to clarify how shape recognition of transparent objects differs from that of opaque objects.

In the first part, we examine several regularities that are potentially related to the shape of transparent objects. For the most part, we restrict our analysis to single images, that is to static situations in which objects, illuminations, and observers are stationary. While some cues remain the same (contour) or exist in modified form (mirror images) as in the opaque case, others are no longer available (e.g. shading, texture). On the other hand, there exist potential cues that are specific to transparent objects. These include background distortions caused by refraction and changes in chromaticity and brightness due to internal absorption (see Figure 4.1b). These regularities are not uniquely related to shape but depend also on many other properties of the object and the scene. While this is a common problem of most visual cues, it is especially pronounced in the case of transparent objects. From a theoretical point of view, it therefore appears likely that the usability of these potential cues depends more heavily on the specific situation than with opaque objects and that the shape recognition differs significantly in both material classes. Since the potential shape cues depend in complex ways on numerous properties of transparent materials, it does not seem appropriate to analyze individual image regularities in isolation. Instead, understanding shape perception in the transparent case appears to require a more comprehensive approach. Further it has to be taken into account that some regularities may only be used as shape cues in certain situations, while in others they appear less appropriate for this purpose.

In the second part of this work, we investigate experimentally how well the visual system infers the shape of transparent objects and on which image regularities it relies in this process. In an experiment, we measured the perceived shape of realistic transparent and opaque objects under natural viewing conditions while varying the availability of different potential shape cues. The results indicate that subjects are able to infer the shape of transparent objects, but less accurately than with opaque objects of identical shape. In addition, some of the image regularities had opposite effects on the accuracy of shape perception, depending on whether the transparent objects were massive or hollow. These results provide strong empirical evidence that the shape perception of opaque and transparent objects actually depends on visual processes that are specific for each material class.

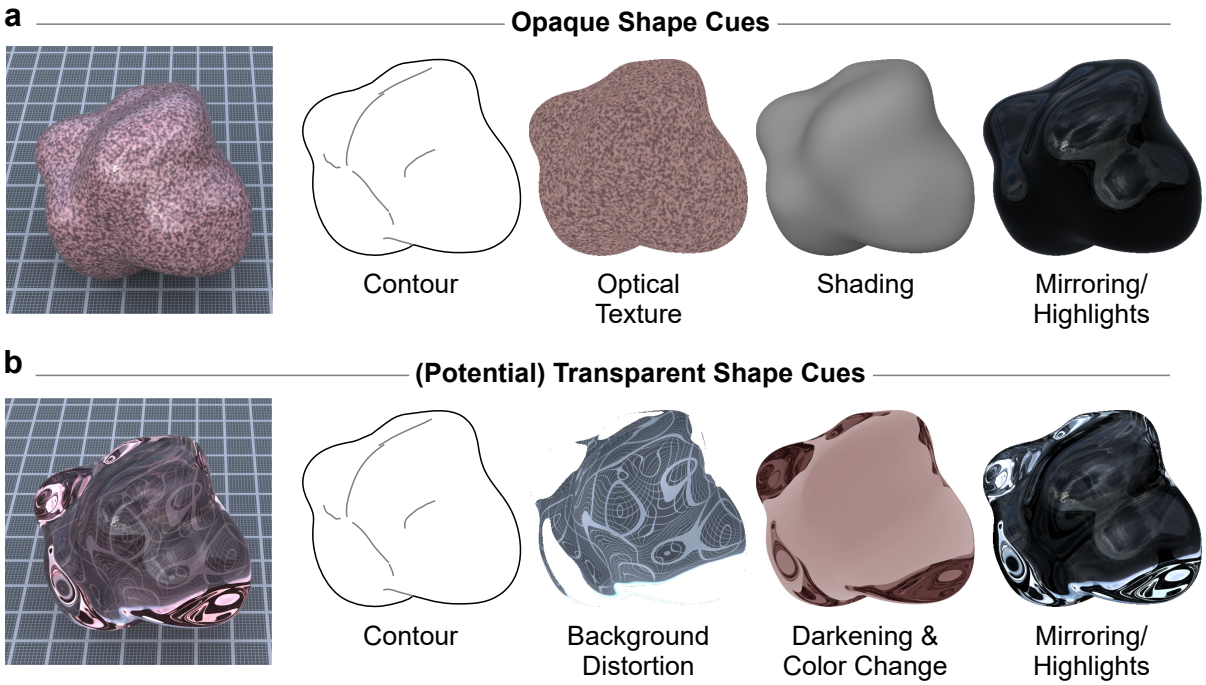

Figure 4.1: Illustration of shape cues for opaque and transparent three-dimensional objects with randomly shaped surfaces (see Appendix 4.A for technical details on this example scene). **(a)** Image regularities that can be used as a cue for the shape of opaque objects include the contour of the object, the density and shape of its texture elements, surface shading, and highlights or mirror images caused by specular reflections. **(b)** For transparent objects some of the regularities known from opaque objects (e.g. shading and texture) are missing, while others remain unchanged (contour) or exist in a similar way (mirror images/highlights). In addition, there are potential shape cues that are specific to transparent objects. These include, for example, background distortions due to light refraction and changes in chromaticity and intensity due to absorption.

4.2 Cues from background distortions due to refraction

Transparent materials generally change the direction of propagation of the light they transmit. This so-called refraction occurs both when light enters and leaves the material. Snell's law describes quantitatively how the degree of refraction depends on the angle of incidence of the light and on the relative optical density of the material and the medium surrounding it. In the image, refraction usually leads to optical distortions of the background visible through a transparent object. This effect is particularly pronounced in massive objects with curved surfaces (cf. Figure 4.1b).

Some previous works already dealt with the question of what role these background distortions play in the perception of transparent materials. For example, Kawabe, Maruya, and Nishida (2015) and Kim and Marlow (2016) could show that certain aspects of these distortions help to distinguish between transparent and opaque materials. Fleming, Jäkel, and Maloney (2011b) suggested that background distortions can serve as a specific cue for the refractive properties of the material (but see Schlüter & Faul, 2014, 2016).

In this work, we focus on the question whether background distortions may also indicate the shape of transparent materials. This problem has so far been investigated mainly in machine vision (Ben-Ezra & Nayar, 2003; Hata, Saitoh, Kumamura, & Kaida, 1996; Morris & Kutulakos, 2011; Murase, 1990, 1992). However, the corresponding findings cannot easily be generalized to the human visual system, because often highly artificial observation situations are assumed and certain scene parameters that are hidden to a human observer are assumed to be known.

To derive the relationship between background distortions and shape, we will consider a simplified situation in which an observer looks through a single slightly waved refracting two-dimensional surface on a flat background (see Figure 4.2a). We further assume an orthographic projection, because in this case, the differences in the degree of refraction of the light that reaches the observer from different surface locations can uniquely be attributed to differences in the local curvature of the surface. With perspective projection, the situation is more ambiguous, because the angle of incidence and thus the degree of refraction does not only depend on the surface orientation, but also on the viewing angle, which varies across the image. In the retinal image, spatial changes in the degree of refraction lead to optical magnifications and compressions of the background texture that is seen through the transparent surface (see Figure 4.2b; since compressions are negative magnifications we will sometimes refer to both simply as magnifications). These magnifications are directly related to the local curvature of the refracting surface (see Figure 4.2c). Convex (i.e. positively curved) areas of the surface generate optical magnifications, concave (i.e. negatively curved) areas produce optical compressions (see Figure 4.2d).

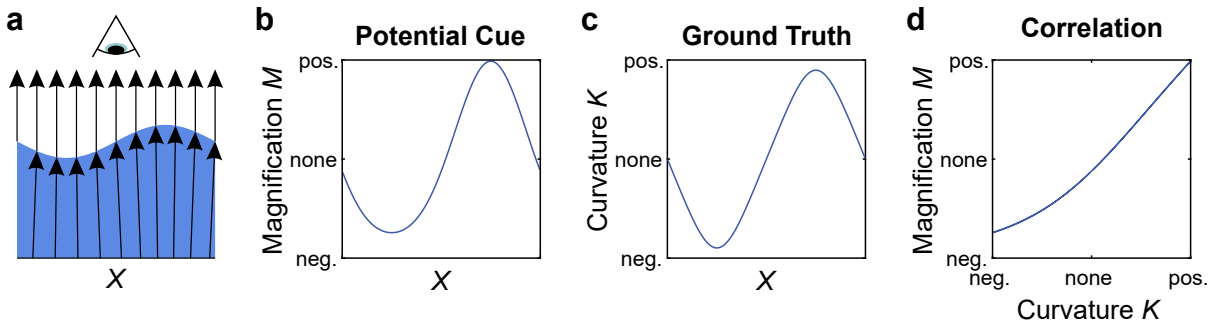

Figure 4.2: Relationship between background distortions and shape. **(a)** Simulation of the light paths reaching an observer that looks at a two-dimensional light-refracting water surface ($R = 1.33$) surrounded by air ($R \approx 1$). An orthographic projection is assumed. Due to the curvature of the refracting surface, its local orientation and thus also the strength of refraction varies continuously. As a result, the points on the underground from which the light rays are reflected towards the observer are unevenly distributed. **(b)** The varying density of the reflection points on the underground causes optical magnifications and compressions in the image. If their density is higher than without refraction, the underground gets optically enlarged ($M > 0$; in this example on the right side), if it is smaller, the underground gets optically compressed ($M < 0$; in this example on the left side). **(c)** Actual curvature K of the surface considered here. At locations where the surface is convex (from the observer's point of view), the curvature is positive, at concave locations it is negative. **(d)** Correlation between magnification in the image (M) and surface curvature (K). Convex surface areas enlarge the underground, concave areas compress it. The more positive the curvature, the greater the optical magnification. The more negative the curvature, the greater the optical compression (i.e. the more negative the optical magnification). The smaller the mean distance between the refracting surface and the underground, the more non-linear this correlation becomes, because the curvature of the surface and its distance to the underground are confounded: For negative curvatures, the distance tends to be smaller. The smaller this distance, the less the influence of the refraction on the image. Thus, magnifications are more pronounced than compressions.

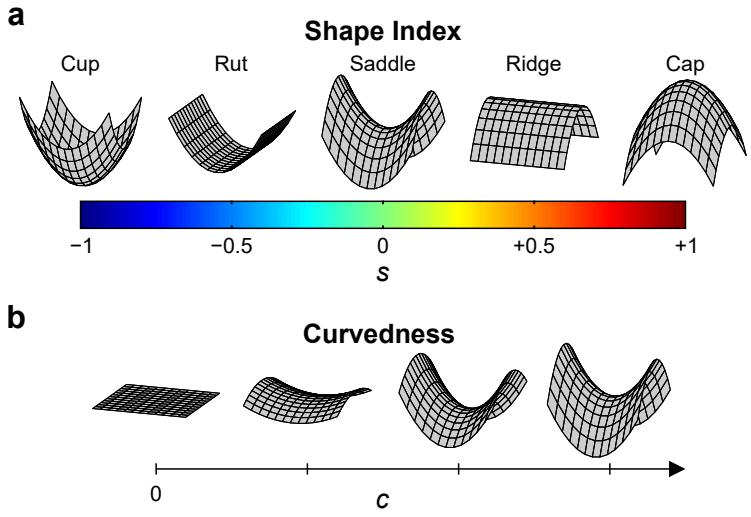

Figure 4.3: Illustration of the shape index and the curvedness, two measures that describe the local shape of a surface. Both were initially proposed by Koenderink and van Doorn (1992) and refer to the local principal curvatures $K_{\min}$ and $K_{\max}$, which fully determine the local shape of a surface (see Appendix 4.B for details). **(a)** The shape index s describes the type of local surface shape on the continuum from spherical concave ($s = -1$, "Cup") to saddle-like ($s = 0$, "Saddle") to spherical convex shapes ($s = +1$, "Cap"). We will continue to use the color scheme used here to refer to different shape index values. The color scale is similar to the one proposed by Koenderink and van Doorn (1992). **(b)** The curvedness c describes the strength of local curvature independently from the type of surface shape (i.e. the shape index s). The higher the curvedness, the more accentuated the surface shape is. In contrast to the shape index, the value of the curvedness depends on the actual size of an object. Here, increasing curvedness values are illustrated with a saddle-like surface ($s = 0$) as an example.

This correlation between distortions and shape persists if the situation is enriched with another spatial dimension. In such a three-dimensional situation, the local curvatures have directional components. For each point of the surface, the curvature is smallest in one direction and largest in another direction, which is orthogonal to the first (principal curvatures $K_{\min}$ and $K_{\max}$). To get an insight how local image distortions and local surface shape are related in the three-dimensional case, a representation of the principal curvatures proposed by Koenderink and van Doorn (1992) seems especially suited. This representation distinguishes between a qualitative and a quantitative aspect captured by the shape index s and the curvedness c, respectively (see Figure 4.3 and Appendix 4.B).

Because a local surface patch is usually curved differently in different directions, the degree of local magnification in general also depends on the directions. For each image location, the magnification is maximum in one direction and minimum in a direction perpendicular to it (for details, see Figure 4.4a). There is an obvious structural similarity between local magnifications and local curvature, which suggests that the two magnification components $M_{\min}$ and $M_{\max}$ are somehow related to the local principal curvatures $K_{\min}$ and $K_{\max}$ of the surface. Figure 4.4b illustrates how specific patterns of minimum and maximum magnifica-

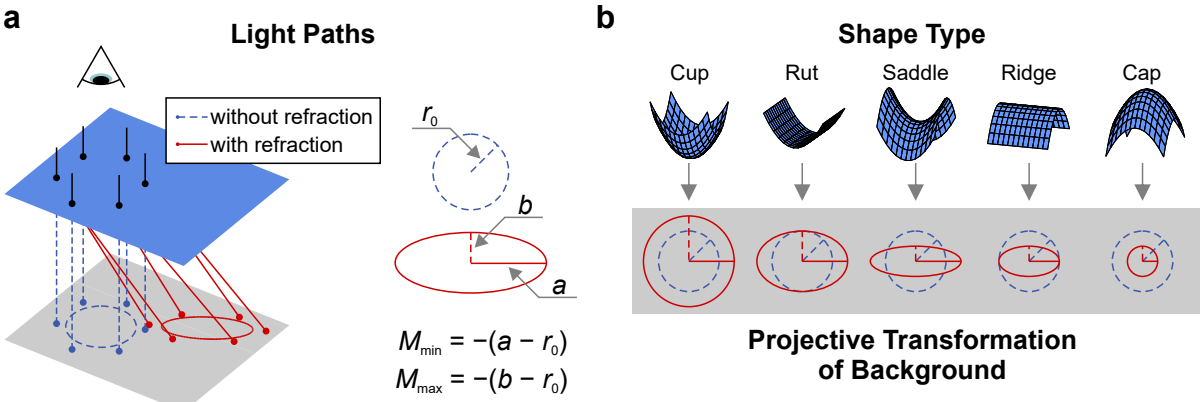

Figure 4.4: Conceptual analysis of the relationship between optical background distortions caused by a light-refracting surface and its shape. **(a)** Illustration of the light paths of six arbitrary light rays reaching an observer in a hexagonal configuration. The geometry of the underground depicted by the undistorted rays can be approximately described by a circle (blue dashed circle). Its radius (r_0) is given by the eigenvalues of the covariance matrix of the reflection points on the underground. The background area depicted by the distorted rays can vary in size, position, and shape. For sufficiently small bundles of light, the form of this background patch can be approximated by an ellipse. Its half-axes a and b are related to the minimum and maximum magnification (M_{min} and M_{max}) with which the ray bundle depicts the underground. More specifically it holds $M_{\text{min}} = -(a - r_0)$ and $M_{\text{max}} = -(b - r_0)$. **(b)** Illustration of how the geometry of an optically distorted background patch (bottom), and thus its minimum and maximum magnifications (M_{min} and M_{max}), are related to the shape type of the refracting surface (top). Like in Figure 4.4a, the blue dashed circles denote the undistorted background patches, while the red circles/ellipses denote the background patches optically distorted by refraction. Specific patterns of minimum and maximum magnifications are related to qualitatively different surface shapes.

tions are related to the local shape type. Based on this relationship, we propose that, given the vector $\vec{M} = (M_{\text{min}}, M_{\text{max}})$, the shape index s can be approximated by the orientation and the curvedness c by the length of this vector $\vec{M}$. More specifically, the angular range $[(1/4)\pi, (5/4)\pi]$ of $\vec{M}$ is mapped to the range $[1, -1]$ of the shape index s. Figure 4.5 depicts the results of a simulation, which confirms the approximate validity of this relationship. Since shape index and curvedness are just an alternative representation of the principal curvatures K_{min} and K_{max} of the surface, the magnifications can also be used to estimate the principal curvatures.

4.2.1 Numerical experiment: Estimating shape from background distortions due to refraction

To analyze how the magnifications caused by a three-dimensional light-refracting surface are related to its principal curvatures and thus its shape, we conducted a numerical experiment, in which the interaction of light rays with a three-dimensional water surface was simulated. Estimates of the local surface shape and the strength of curvature were then derived from the

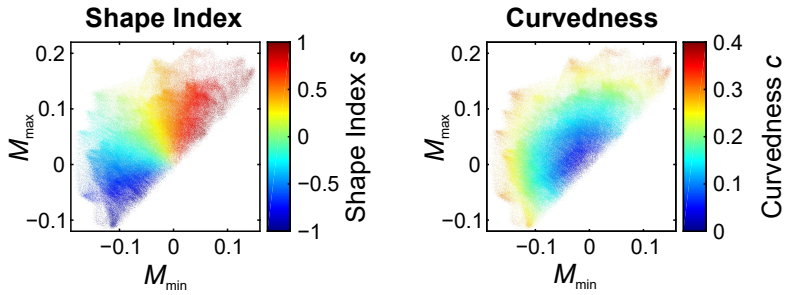

Figure 4.5: Simulation performed to verify the relationship between the minimum and maximum magnifications (M_{min} and M_{max}) and the shape index s (left) and the curvedness c (right). The results are based on a large number of bundles of light (cf. Figure 4.4a) passing through a slightly curved water surface like the one in Figure 4.6a. The results show that there is indeed a close relationship between the magnifications and the shape: The orientation of the vector $\vec{M} = (M_{\text{min}}, M_{\text{max}})$ approximates the shape index s and its length approximates the curvedness c.

optical distortions of the surface beneath the water. To test how robust these estimates are, we varied both the amplitude of the water surface as well as its distance to the underground.

4.2.1.1 Stimuli

The stimuli were computer-generated images of a water surface that was modeled with the 3D modeling software Blender (Blender Foundation, 2015). The surface was created from a plane of size 140.4×140.4 mm, which consisted of 24200 triangular faces. The vertices of this plane were displaced along their normal direction (modifier "Displace" with parameters "Direction" = "Normal", "Texture Coordinates" = "Local", "Midlevel" = 0.5, and "Strength" = 0.5). The amount of displacement was determined by the intensity of a Perlin noise texture (texture "Cloud" with parameters "Noise Basis" = "Original Perlin", "Size" = 1, and "Depth" = 0 and options "Grayscale" and "Soft" selected). The resulting maximum amplitude of the water surface was $a = 16.9$ mm. The mean distance to the underground was set to $d = 70.7$ mm.

In two additional conditions, either the amplitude of the water was adjusted to $4a$ by adjusting the "Size" parameter of the Perlin noise texture accordingly or the distance to the underground was set to $8d$.

An orthographic camera was located above the water surface. The scene was rendered as a high dynamic range image with the physically based Cycles renderer (Blender Foundation, 2015; image size 520×520 px, 2048 samples/px, color depth 16 bit/channel). The material properties of the wavy surface have been adjusted to match those of water. However, to isolate the effects of optical distortions, Fresnel effects and specular reflections have been ignored (material shader "Refraction BSDF" with parameters "Distribution" = "Sharp", "Color" = RGB(1, 1, 1), "Roughness" = 0, and "IOR" = 1.33).

4.2.1.2 Procedure

To gain information about the actual light paths, the underground was provided with color gradients, so that each point of the underground had a unique color. Using the color of a pixel in the rendered image as an index, it was therefore possible to determine the point on the underground that it depicts. To determine the degree of displacement due to refraction, the same scene was rendered with and without the water surface. Then, the positions of pixels of same color in both rendered images were compared. In this way, both the light paths (beginning with the corresponding points on the underground) and the displacement due to refraction (given by the distance of corresponding points in the images) could be calculated.

Based on this procedure, the minimum and maximum magnifications (M_{min} and M_{max}) of a total of 492246 evenly distributed bundles of light B_i passing through the water surface were calculated. Each bundle of light consisted of six arbitrary light rays arranged hexagonally (cf. Figure 4.4a). The distance between two adjacent rays was 0.27 mm. The orientation and length of the vector $\vec{M}_i = (M_{\text{min},i}, M_{\text{max},i})$, which denotes the magnifications for bundle of light B_i, was then used to estimate the corresponding shape index and the curvedness, respectively. The veridical principal curvatures were calculated using a procedure proposed by Rusinkiewicz (2004). From them the veridical shape index and curvedness were also calculated.

It should be noted that the procedure used here is based on the inverse of the real image generation process where light propagates after a diffuse reflection at the underground in the direction of the observer. However, this has no influence on the distortions of the background in the retinal image. This fact is exploited in the path tracing algorithm used in much render software that also starts form the viewing point and follows the light rays backwards to surfaces and light sources in the scene.

4.2.1.3 Results

The stimuli and results of the numerical experiment are shown in Figure 4.6.

Slightly Curved Surface The leftmost two diagrams in Figure 4.6a show a subset of the light paths simulated for the slightly curved water surface. Both the minimum and maximum components of the curvature are strongly correlated with the minimum and maximum components of the magnification (see Figure 4.6a third diagram from right). Estimating the shape index and curvedness from the magnification components leads only to minor deviations from ground truth (see Figure 4.6a rightmost two diagrams). It should be noted, however, that estimated and veridical curvedness values can differ from each other by an a priori unknown factor. This factor depends on arbitrary parameters, such as the diameter of the simulated bundle of light. Estimating the absolute strength of curvature from the magnifications would therefore require an appropriate anchoring.

Strongly Curved Surface If the amplitude of the water surface is quadrupled, the higher curvature causes the light rays to converge and diverge more strongly than with the less wavy water surface (see Figure 4.6b leftmost two diagrams). As a consequence, the rays from the underground to the water surface may cross. This happens whenever the positive curvature of a surface area exceeds a certain value. Below this critical value, an increasing curvature is accompanied by an increasing magnification. At the critical curvature the maximum possible magnification is reached and all light rays of the corresponding image area are converging to a single point on the background. If the curvature is further increased beyond this critical value then corresponding points on the underground start to move away from each other and the magnification begins to decrease again (we will refer to this as "magnification inversion"). At extreme curvatures the magnification can even become negative and turn into a compression. If a magnification inversion occurs, the correlation between magnification and curvature is no longer monotonous, but inversely v-shaped (see Figure 4.6b third diagram from right). If shape type and curvature strength are estimated without taking the magnification inversion into account, this results in larger deviations from the veridical values (see Figure 4.6b rightmost two diagrams).

Increased Background Distance A magnification inversion can also occur, if the distance to the underground is increased (see Figure 4.6c leftmost two diagrams). Although the direction changes at the water surface are relatively small, the rays can nevertheless cross each other because of their greater travel distance. Correspondingly, the correlation between curvature and magnification is no longer monotonous (see Figure 4.6c third diagram from right) and there are larger deviations between the estimated and veridical shape parameters (see Figure 4.6c rightmost two diagrams). In this case, however, the strength of curvature can be estimated more accurately than the shape type.

4.2.1.4 Conclusion

The results of the numerical experiment show that in certain situations, distortions of the underground can be used to estimate the shape of the water surface that lies above it. However, our analysis has also revealed that an approximately linear relationship between shape and background distortions can only be assumed in cases where light rays do not intersect and only "simple" optical compressions and magnifications occur.

4.2.2 Generalization and open questions

Estimating the shape of transparent objects from background distortions caused by light refraction is a very complex problem. The approach described in the foregoing, which explores the usability of distortions as a shape cue on a rather fundamental level, can therefore only be

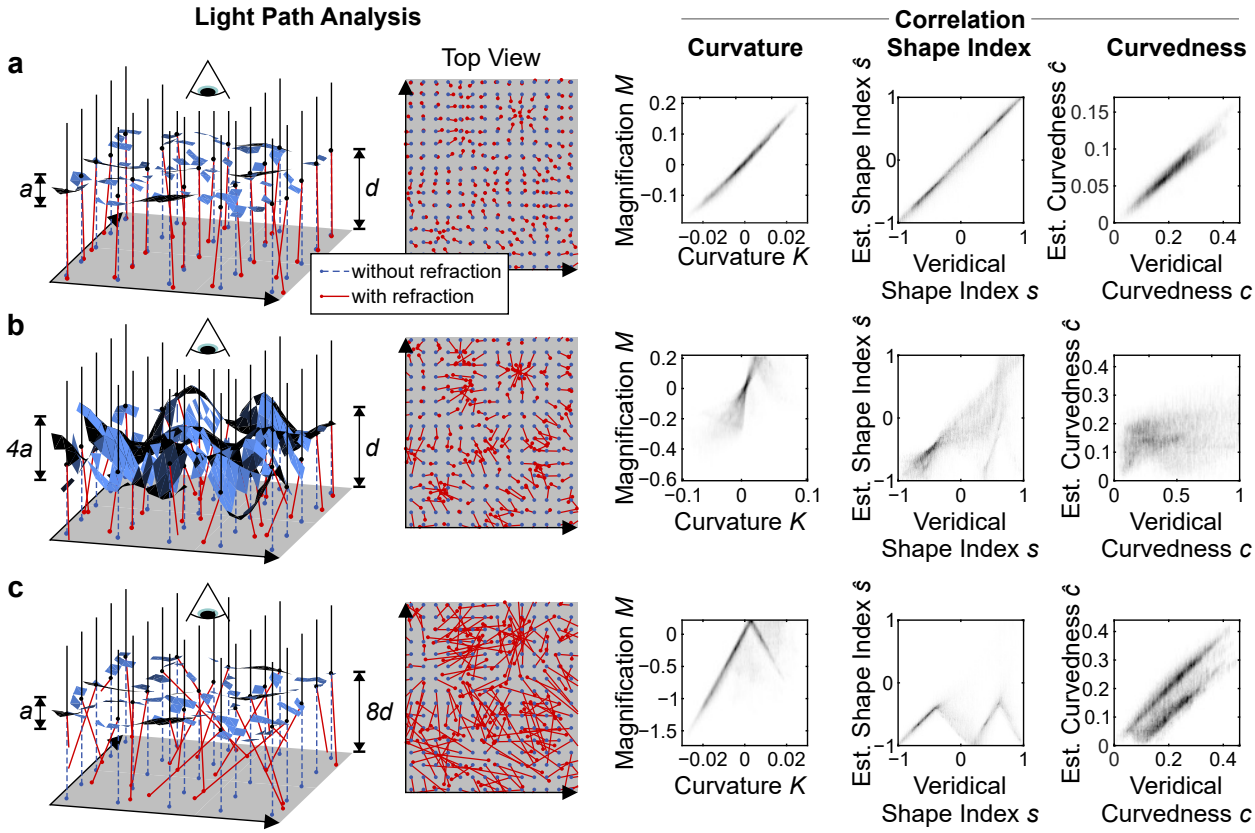

Figure 4.6: Results of the numerical experiment. The leftmost two column show a subset of the simulated light paths (the mesh of the water surface is shown in reduced resolution). The three rightmost columns show the correlation between estimated and veridical shape in terms of magnification/curvature (minimum and maximum components are considered simultaneously), shape index and curvedness. **(a)** Results for the slightly waved water surface. Estimated and veridical shape parameters correspond well. **(b)** Results for the strongly waved water surface. Some light rays cross each other in such a way that there is a magnification inversion. As a consequence, optical magnifications are no longer unambiguously related to local curvature. **(c)** Results for the slightly waved water surface with a greater distance to the underground. The magnification inversion is even more pronounced than in the previous condition, so that the correlation between estimated and veridical shape type is alternately positive and negative. The correlation between estimated and actual curvedness is characterized by two branches running parallel to each other, whose offset results from the fact that here, magnification inversions occur only for curvatures $K > 0.004$.

a starting point for more thorough investigations. In the following we will discuss a number of topics that in our view need to be addressed in further research. First, we address the question of how the veridical optical magnifications, which are not directly accessible to an observer, could be estimated from the image, and to what extent orientation maps could help with this. Then we address the point that under natural viewing conditions, optical distortions do not necessarily indicate the intrinsic local curvature of a surface, but its curvature relative to the observer, that is, the rate at which the surface orientation changes for successive locations in the image.

Estimation of distortions from the image Under normal viewing conditions, the veridical optical magnifications are not directly accessible to an observer. The question therefore arises, whether optical distortions can be estimated from the retinal image alone. Provided that the background has sufficient structure and contrast to reflect the optical distortions, they are visible in the image as direction-dependent variations in texture density, where optical magnifications tend to produce lower texture densities than optical compressions (see Figure 4.7a). The minimum and maximum texture densities $D_{\min}$ and $D_{\max}$ in a small region around a specific location in the image could therefore be used as estimates of the corresponding maximum and minimum optical magnification $M_{\max}$ and $M_{\min}$, respectively (see Figure 4.7b). As it was shown above, the latter are closely related to the shape index s and the curvedness c, which in turn uniquely determine the maximum and minimum principal curvatures $K_{\max}$ and $K_{\min}$.

However, inferring optical magnifications from texture densities is difficult for several reasons. First, texture density variations in the image cannot always be attributed to differences in optical magnification. They may also be caused by properties of the background itself (e.g. variations of its intrinsic texture density). Reliable inferences therefore require that the background is either sufficiently regular or that further information about the undistorted background is available. Second, it is obvious that a certain texture density in the image can only be interpreted as a magnification or a compression if the texture density corresponding to the undistorted state is known. If parts of the undistorted background are directly visible at another point of the image or at another point in time, this information could be used as a reference (cf. Fleming et al., 2011b), provided that the background is spatially and temporally sufficiently homogeneous. If the background is entirely located behind the transparent material, as in the case of the water surface analyzed above, one can only refer to image statistics. One could, for example, estimate the undistorted background texture density by the mean value of minimum and maximum texture densities across the image.

For the refracting surface used in the numerical experiment (cf. Section 4.2.1), we analyzed exemplarily how well the undistorted initial state can be estimated from the distorted state. For this purpose, we compared the semi-axes a_i and b_i of the n distorted background areas

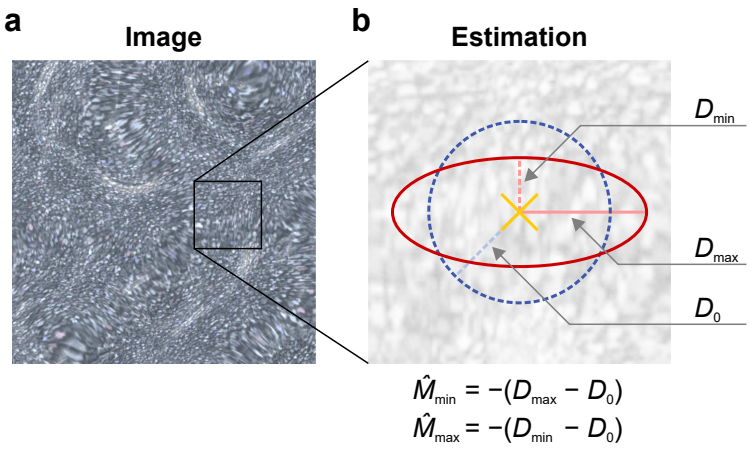

Figure 4.7: Relationship between directional variations of texture density, light path geometry and local shape. **(a)** Image of the water surface simulated in the numerical experiment (cf. Section 4.2.1) from the observer's point of view. The optical magnifications and compressions caused by light refraction are visible as direction-dependent variations in the texture density of the underground visible through the water. In addition, slight reflections of the environment are visible on the water surface. **(b)** A possible estimation of the directional magnification from the directional texture density. The texture density at one point of the image (yellow cross) is always maximum in a certain direction ($D_{\max}$, horizontal red line) and minimum in a direction orthogonal to it ($D_{\min}$, vertical dashed light red line). For regular background textures, $D_{\min}$ and $D_{\max}$ behave proportionally to the minimum and maximum expansion b and a of the background area depicted at that image position (overlayed as red ellipse, cf. Figure 4.4a). In order to get estimates $\hat{M}_{\min}$ and $\hat{M}_{\max}$ for the absolute magnifications $M_{\min}$ and $M_{\max}$, which are related to the local principal curvatures $K_{\min}$ and $K_{\max}$, $D_{\min}$ and $D_{\max}$ would have to be compared with the original (i.e. undistorted) texture density D_0 (dashed light blue line). For regular backgrounds, this is proportional to the size of the undistorted area of the background depicted at the image location (overlayed as dashed blue circle, cf. Figure 4.4a). One problem with estimating the shape from the directional texture density arises from the fact that D_0 is not directly available in the distorted image (see text).

i not with the actual radius of the undistorted background area (r_0) but with the mean of all semi-axes ($\hat{r}_0 = \text{mean}(a_1, \ldots, a_n, b_1, \ldots, b_n)$). The correlation between minimum and maximum magnifications estimated in this way and the minimum and maximum curvatures remains virtually the same. Thus, at least in the situation considered here, estimating optical magnifications solely from the distorted background area appears quite plausible. In other situations, however, the described procedure might be less suited. For example, a surface with mainly positive curvature (e.g. a drop of water on the ground) will on average magnify the underground. In such a case, the undistorted texture density would tend to be underestimated.

Orientation maps as a potential cue Chen and Allison (2013) suggested that the shape of transparent objects could be estimated based on the orientation maps of the background distortions caused by them. Orientation maps were initially proposed by Fleming, Torralba, and Adelson (2004) as a shape cue for opaque objects (see also Fleming, Holtmann-Rice, & Bülthoff, 2011a). An orientation map describes spatial image structures (irrespective of their underlying cause) by indicating at each pixel the dominant orientations of the texture (i.e. in the notation used in the foregoing, the direction of D_{min} and D_{max}) and the direction-related strength of the texture density (i.e. the relative size of the texture density in different directions). Chen and Allison (2013) surmised that orientation maps could serve as a shape cue also in the transparent case, because the image distortions caused by transparent and specularly reflecting opaque objects are similarly related to object shape. In both cases larger surface curvatures cause larger distortions in the mapping of the environment/the background to the image. At present, it is an open empirical question, whether orientation maps actually play a role in the perception of transparent objects. However, there is some reason to be skeptical, because according to the derivation outlined above, orientation maps do not contain sufficient information for a successful estimation of object shape: While the dominant orientation of the texture at a certain image point indicates the orientation of the minimal curvature, information about the sign and absolute size of the principal curvatures is lacking. The direction-dependent strength of the texture, which is estimated from the minimum and maximum local filter responses, only provides information about their relative size and this is not sufficient to estimate the shape type (i.e. the shape index) of a surface. This means, for example, that convex surface areas cannot be distinguished from concave ones.

Intrinsic vs. observer-dependent shape So far, we have assumed that optical distortions depend *solely* on the local curvature of a surface. However, under natural viewing conditions this is generally not the case. If the observed surface is not orientated perpendicular to the viewing direction, factors related to the optical projection may also play an important role. As a consequence, magnifications in the image may be less correlated with the intrinsic curvature of a surface than with its curvature relative to the observer, that is, with the rate at which

the surface orientation changes "in the image" (see Figure 4.8). At places where the surface is strongly inclined to the viewing direction, for example near the rim of an object, it is therefore difficult to infer the intrinsic curvature of the surface from magnifications in the image. A similar issue has been highlighted by Fleming et al. (2004) in the case of mirror objects. It is not yet clear what the consequences of this dependency on the observer are and how an estimate of the intrinsic, observer-independent shape of a surface can be gained from view-point dependent information.

Use as a cue in more complex situations In the transparent case, the light transport can in principle be of any complexity. For example, in a typical transparent object with a closed outer surface, light reaching the observer from the background is usually not only refracted once but at least twice. The question arises how background distortions are related to shape in such cases. An exemplary analysis of the corresponding light paths shows how complex the depiction of the underground and thus the distortions can be in such cases (see Figure 4.9a). If the light is refracted twice, the distortion at each image point is simultaneously determined by the shape of two different parts of the object's surface. Inferring the shape of *one* of the surfaces separately is difficult, if not impossible, because the distortions in the image are not merely the sum of two individual distortions that are independent of each other. If background distortions can contribute at all to the perception of shape in such situations, it seems at least necessary to make additional assumptions about the range of possible shapes that may occur in a given context.

In a way, the situation becomes even more complex when light reaching the observer is refracted more than twice as for example in the case of hollow objects. However, the larger number of refractions does not necessarily lead to stronger distortions. With relatively small wall thicknesses, the distortions can be much weaker than for massive objects of the same shape (see Figure 4.9b). Although the fundamental problem of inferring the shape of *one* of the surfaces that contribute to the image distortion remains the same, there are two favorable circumstances that may facilitate the estimation of object shape. First, the correlation between distortions and curvature is hardly affected by magnification inversions as most light rays that reach the observer do not cross. Secondly, two of the four surfaces involved in the distortion have often a very similar shape (outer and inner front and back of the object). Their influence on the distortion is therefore relatively similar and could thus serve as a joint cue for the shape of each pair of surfaces.

Further complexity arises from the fact that in the general case the background behind the distorting material does not have to be flat but can consist of arbitrary objects. The distortions then also depend heavily on the properties of the background scene. This includes not only the shape and material of objects seen in the background, but also their position relative to

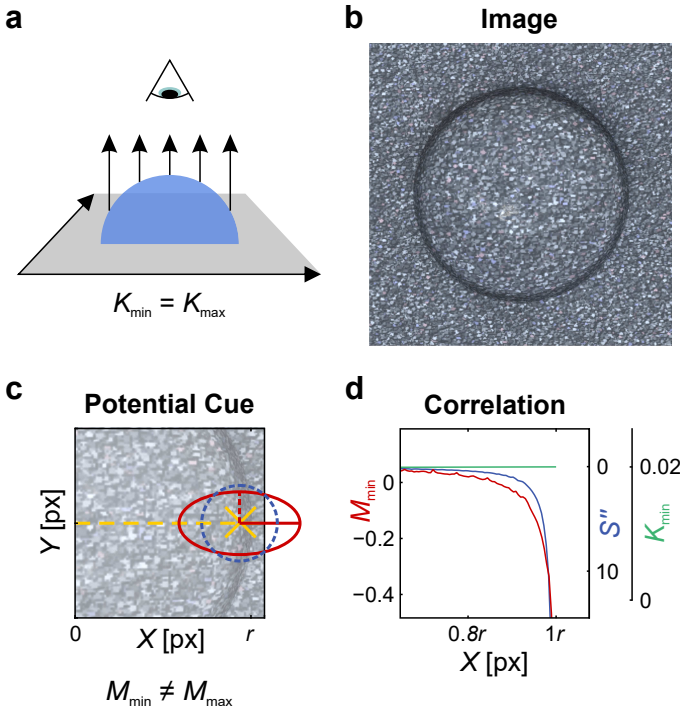

Figure 4.8: Example of a case where background distortions are not directly related to the intrinsic curvature of a surface. **(a)** This example shows a hemispherical drop of water ($R = 1.33$, radius $r = 50$ mm) placed on a surface and surrounded by air ($R \approx 1$). At each point of its curved surface its intrinsic principal curvatures are identical and constant ($K_{\min} = K_{\max} = 1/r = 0.02$). Accordingly, the shape index is constant ($s = 1$). **(b)** Image of the water drop and optical distortions of the underground from the observer's point of view. In contrast to its principal curvatures, the optical magnifications created by the drop are not the same at every point. Especially near the rim of the drop, the directional dependence of the optical magnifications is increased. **(c)** Analogous to Figure 4.4a, the true geometry with which the background is depicted is shown for a spot close to the rim of the drop (yellow cross). In the horizontal direction there is an optical compression (i.e. a negative optical magnification), in the vertical direction there is a (weak) magnification (i.e. $M_{\min} \neq M_{\max}$). **(d)** Correlation of the magnification with different curvature measures. Here, the horizontal component of the optical magnification (in this case $M_{\min}$, red line), the intrinsic curvature (with $K_{\min} = K_{\max}$, green line), and the curvature relative to the observer, which is the second derivative of the surface function in the image space (S'', blue line) are shown for the central axis of the drop (dashed yellow line in Figure 4.8c). Especially near the drop's rim the course of $M_{\min}$ resembles that of the curvature relative to the observer (S'') rather than that of the intrinsic curvature $K_{\min}$.

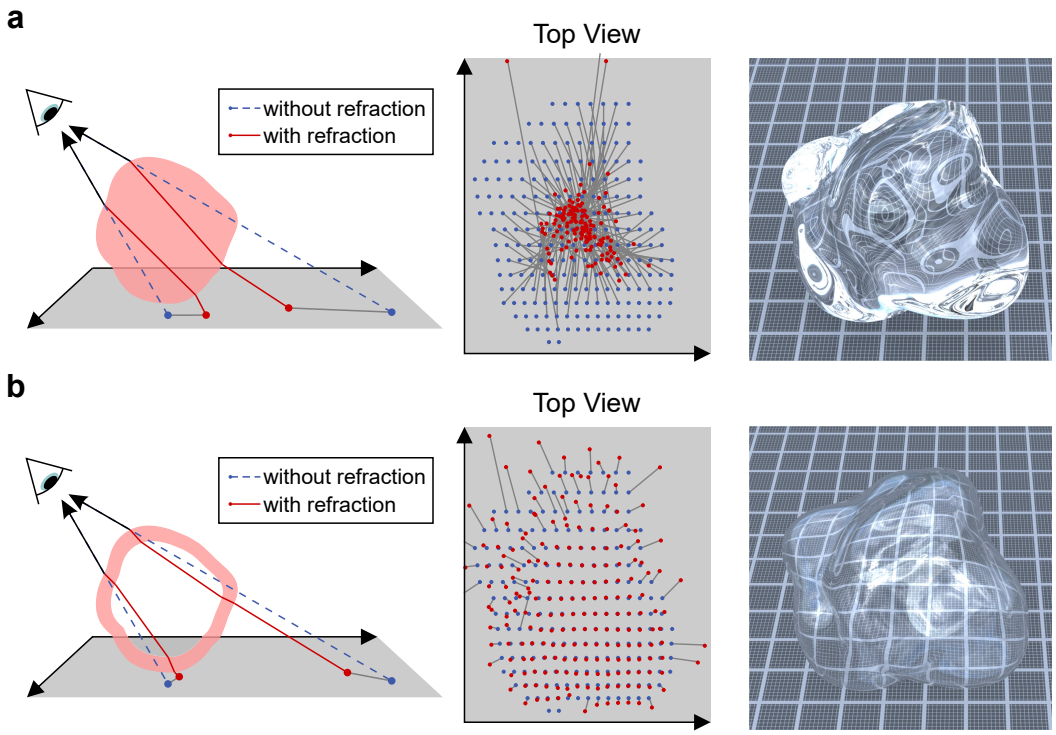

Figure 4.9: Light path simulations for a massive and a hollow three-dimensional transparent object similar to the one shown in Figure 4.1b. Note that the respective diagrams in the center show the results of actual simulations, while the left diagrams show schematic illustrations. The simulations were performed similar to the procedure described in the numerical experiment (cf. Section 4.2.1). However, a perspective projection was used. See Appendix 4.A for technical details on the images shown on the right. **(a)** On average, the massive object magnifies the underground. At some places, light rays cross each other in such a way that there is a magnification inversion. In addition, due to total reflections within the object, some of the light rays that reach the observer are never reflected by the underground but originate directly from other elements of the scene (blue dots without red partners). The distortions of the background caused by the varying directional magnifications are visible in the image as directional variations of the texture density (image on the right). At locations where total reflections occur within the object, the background is either not visible at all or at least not without further reflections and/or light refractions (here, this is mainly the case near the rim of the object). **(b)** Although hollow objects refract light more often, their distortions can be much weaker, if their wall thickness is relatively small and does not vary too much. Here, displacements of the reflection points are substantially smaller than for the massive object and so are the direction-dependent variations of the texture density in the image (image on the right).

the transparent object. Using distortions as a shape cue in such situations would be even more difficult.

In addition to the aspects considered so far, temporal changes in the scene can also influence background distortions. Although this dynamic further increases the complexity of the scene, it may also provide additional shape cues, as is suggested by works from machine vision (Morris & Kutulakos, 2011; Murase, 1990, 1992). In addition, dynamics could help to identify individual shape cues in the image that are confounded in a static stimulus. For example, the contributions of the distorted background and the distorted mirror images might be more easily separated, because the image features related to these causes would in general move differently in the image.

4.3 Cues from changes in intensity and chromaticity due to absorption

Many transparent materials absorb parts of the transmitted light. Absorption is described by the Bouguer-Lambert-Beer law and it can change both the intensity of the light and its spectral distribution. In the image, this leads to a darkening and to changes in chromaticity. Since the strength of these effects depends on the length of the light path inside the object, they are related to the thickness of an object and may thus indirectly contribute to the recognition of the object's shape. For simple transparent filters, Faul and Ekroll (2011) have already shown that estimating the thickness from transmission and saturation can succeed in certain situations. In this section, we will investigate to what extent these findings can be generalized to objects of more complex shape and to more natural viewing conditions. Since estimates based on changes in luminance have proven to be much more robust than estimates based on saturation changes (Faul & Ekroll, 2011), we will focus on the role of absorption-related darkening.

Darkening due to absorption can only be used as a shape cue if it can be discerned from other causes of darkening in the image, such as spatially varying background reflectance properties or changes in transmittance and reflectance due to Fresnel effects. Comparable problems arise in identifying the shading of opaque objects, because this requires to isolate the darkening in the image due to orientation changes of the surface from other causes of darkening, and it is quite possible that similar perceptual mechanisms are used in both cases.

A further problem arises from the fact that the relationship between darkening and thickness is significantly influenced by the absorption properties of the material and by factors such as light refraction and total reflection. Objects of different thicknesses can therefore lead to similar absorption-induced darkening, if the absorption coefficients of their materials are properly chosen. This means that the *absolute* thickness of a material can only be estimated if the absorption coefficient is known. Without this knowledge, at most information about

the *relative* thickness of can be obtained. Anchoring such relative thickness information would then require an integration with other shape cues. The situation becomes particularly complex when the absorption properties of the material are spatially inhomogeneous. In such cases, even relative thickness information can no longer be deduced unambiguously from the darkening. In addition, due to refraction and/or total reflections inside the object, it cannot be assumed that the light rays reaching the observer have crossed the object along a straight line in the respective viewing direction. As a consequence, the distance traveled by the light within the material is generally not identical to the thickness of the object along the respective line of sight.

4.3.1 Numerical experiment: Estimating shape from intensity changes due to absorption

To analyze the correlation between absorption-induced darkening and object thickness and to test how robust thickness values estimated from the darkening are against influences from refraction and total reflection, we performed a numerical experiment.

4.3.1.1 Stimuli

The stimulus material consisted of computer-generated images of 15 randomly shaped transparent objects. The objects were created with the 3D modeling software Blender (Blender Foundation, 2015). The object mesh was based on a icosahedron whose faces were subdivided six times. The resulting icosphere consisted of 81920 triangular faces and was adjusted to a diameter of 100 mm. The icosphere was deformed by translating its vertices along their normal direction (modifier “Displace” with parameters “Direction” = “Normal”, “Midlevel” = 0.5, and “Strength” = 1). The amount of displacement was determined by the intensity of three-dimensional Perlin noise (texture “Cloud” with parameters “Noise Basis” = “Original Perlin”, “Size” = 1, and “Depth” = 0 and options “Grayscale” and “Soft” selected). To gain different shapes, the noise was probed at different locations.

A perspective camera was placed at a distance of 400 mm to the center of the object (vertical field of view 44.10°). The objects were rendered as high dynamic range images with the Cycles renderer (Blender Foundation, 2015; image size 1040×1040 px, 2048 samples/px, color depth 16 bit/channel). An infinitely distant spherical emitter was set to homogeneous white (surface shader “Background” with parameters “Color” = RGB(1, 1, 1) and “Strength” = 1).

4.3.1.2 Procedure

Each of the 15 objects was rendered three times with different material properties to gain information about the veridical object thickness, absorption-induced darkening and areas of total reflection.

In the first run, object and scene parameters were chosen so that the image contained information about the veridical object thickness. To this end, the object was defined not to have a distinct surface, but to consist exclusively of a homogeneous absorptive material (no surface shader; volume shader "Volume Absorption" with parameters "Color" = RGB(.8, .8, .8) and "Density" = 0.03). The resulting image shows the decrease in light intensity caused by absorption. The intensity values of this image were transformed to represent the actual object thickness along the respective viewing directions. To this end, each pixel of the grayscale image p_{xy} was transformed so that $p'_{xy} = -1/a \times \log(p_{xy}/1)$, where a corresponds to the absorption coefficient that was defined during the rendering. With $C = 0.8$ and $D = 0.03$ corresponding to the values of the parameters "Color" and "Density" used by the Cycles renderer, a is given as $a = (1 - C) \times D$.

In the second run, object and scene parameters were chosen so that the image represents absorption-induced darkening. To this end, a surface shader has been added to the object material that takes refraction and total reflections into account but ignores non-total specular reflections and Fresnel effects (see Appendix 4.C for the Blender node setup that was used to define the Cycles material). The intensity values in the image then corresponded to the amount of darkening. This can be described as the ratio $I(d, \lambda)/I_0(\lambda)$, where $I_0(\lambda)$ denotes the intensity of light of wavelength λ before it enters the material and $I(d, \lambda)$ the intensity of the exiting light after it travels a distance d inside the material. Based on this image, local estimates of the object thickness were calculated. Due to the exponential relationship between darkening and light path length inside the object, the local thickness of the objects was estimated by $\hat{t} = -1/a(\lambda) \times \log(I(d, \lambda)/I_0(\lambda))$. This can be considered as the estimate of an *ideal observer*, who can unambiguously identify the darkening due to absorption and knows the absorption coefficient of the material.

In the third run, object and scene parameters were chosen so that image areas with total reflection can be identified. To this end, the object material was changed so that it refracts light (without taking Fresnel effects into account) but does not reflect or absorb it (surface shader "Refraction BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(1, 1, 1), "Roughness" = 0, "IOR" = 1.49; no volume shader). With this setup, all pixels of the image hit by light rays that were totally reflected within the object at least once along their path were displayed in black. All other pixels were white. This black and white image was then used as a mask to consider areas with or without total reflections separately.

4.3.1.3 Results

Figure 4.10 summarizes the simulation results. While both refraction and total internal reflection reduce the correlation between darkening and object thickness, the negative effect is much weaker for light that is only refracted than for light that is also subject to total reflection (see Figure 4.10a). The errors that occur if the object thickness is estimated without taking

these negative influences into account vary across the surface. Near the rim of the objects, where total reflections mainly occur, the errors are much larger than in the middle of the object area, where light gets mainly only refracted (see Figures 4.10b and 4.10c).

4.3.1.4 Conclusion

The results of the numerical experiment show that absorption-induced darkening can provide information about the thickness of a transparent object even without explicitly taking refraction into account. However, as soon as total reflections occur within the object, the darkening hardly provides any useful information about object thickness. An appropriate strategy could therefore be to use darkening only as a thickness cue in image regions without total reflections. This assumes the ability to detect areas affected by total reflections. Typical image properties related to total reflections that might be used for this purpose include strong darkening, bright reflections, and high saturation.

4.3.2 Generalization and open questions

Although absorption-related shape cues seem less complex than the distortion related ones discussed above, there are still many open questions that need to be addressed in future research. In the following, we discuss whether absorption-induced darkening can provide shape information in cases where a transparent object is hollow and how thickness information could be integrated into a specific object shape.

Hollow objects With hollow objects, the correlation between darkening caused by absorption and object thickness is normally considerably weaker, because light that passes through such objects is refracted more often and only some sections of its path inside the outer surface of the object lead through the object's material. The length of these sections can deviate so much from the object thickness that the darkening due to absorption does not provide any useful information about it. Thus, the best strategy seems to be to simply ignore the darkening as a thickness indicator. How hollow objects can be identified is a separate problem. The identification could, for example, be based on very weak background distortion. However, this would only be a rough heuristic because massive objects can also cause weak background distortions, for example if they are rather thin and/or only slightly curved.

Integration of thickness information In order to serve as a shape cue, the thickness information obtained from the darkening in the image would have to be integrated into a specific object shape. This is difficult because the thickness indicated by the darkening is compatible with an arbitrary number of pairs of different front and back surface shapes. For example, thickness information is invariant against an exchange of the front and back surface.

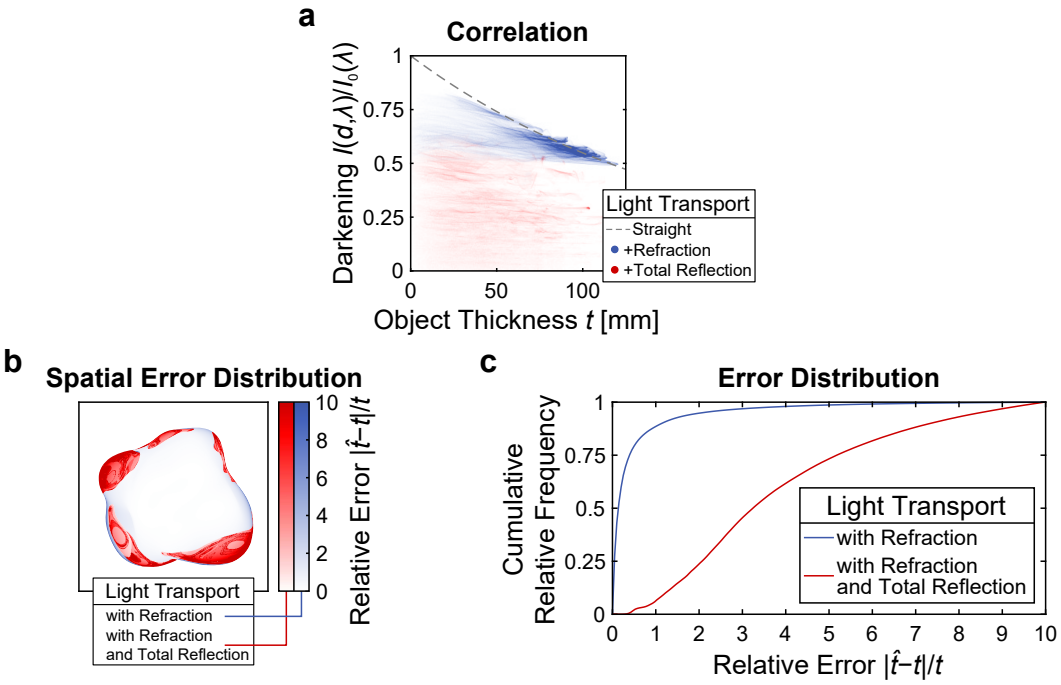

Figure 4.10: Results of the numerical experiment. **(a)** Influence of light refraction and total reflection on the correlation between darkening in the image and object thickness for all 15 simulated objects. The saturation of a point corresponds to the frequency with which a certain combination of darkening and object thickness occurred. The correlation for light that has been totally reflected at least once (red points) is substantially weaker than for light that has only been refracted (blue points). For comparison, the dashed gray line in the plot shows the relationship between darkening and thickness for (hypothetical) light that is neither refracted nor totally reflected. **(b)** Typical spatial error distribution demonstrated at one of the simulated objects. In places where the light path includes total reflection (red areas), the error (represented as saturation) is much greater than where the light was only refracted (blue areas). Since total reflections mainly occur near the object's rim, this is where the errors are largest. The negative influence of refraction is also largest near the rim of the object, where light hits the surface at a more shallow angle. **(c)** Distribution of the error for image areas with (red) and without (blue) total reflections for all 15 simulated objects. In 94 % of the areas with total reflection, the estimated thickness deviates by more than 100 % from the veridical one (i.e. $|\hat{t}-t|/t > 1$). In the area without total reflection this applies to only 11 % of the cases.

These ambiguities could be reduced by general heuristics about object shapes. For example, shapes that are symmetrical in depth may be preferred over asymmetrical shapes. Another way to reduce the space of compatible shapes would be to integrate the thickness information with the information from other shape cues. If these other cues provide sufficient information about the front shape of an object, absorption-induced darkening could then in principle be used to estimate the rear shape of a transparent object.

Furthermore, it is possible that darkening plays additionally (or even exclusively) a "passive" role in the perception of shape, i.e. that it indirectly influences other potential shape cues. For example, more absorptive materials can reduce the visibility of the background and internal mirror images while enhancing the visibility of the reflectance on the front side. However, it is difficult to predict the influence of such interactions on shape perception.

4.4 Cues from mirror images due to specular reflections

Light that hits a transparent material is usually not completely transmitted. Instead, a part of the light is specularly reflected at the material's surface. The relative amount of reflected and transmitted light is described by the Fresnel equations and depends on the angle of incidence and the refractive properties of surround and material. Like in the opaque case, the microstructure of the surface determines how strongly the light scatters around the mirror direction. Here, we focus on ideally smooth surfaces that reflect the incident light solely in the mirror direction. In the retinal image, these ideal specular reflections lead to a sharp mirror image of the environment on the outer surface of the transparent object (cf. Figure 4.1b rightmost).

Opaque and transparent objects of the same shape and surface structure produce geometrically identical mirror images. This suggests that the existing findings on the role of gloss in the perception of opaque objects' shape (Adato, Vasilyev, Ben-Shahar, & Zickler, 2007; Adato, Vasilyev, Zickler, & Ben-Shahar, 2010; Fleming et al., 2004; Muryy, Welchman, Blake, & Fleming, 2013; Oren & Nayar, 1997; Savarese, Chen, & Perona, 2004a, 2005; Savarese, Fei-Fei, & Perona, 2004b; Savarese & Perona, 2001, 2002) can be transferred to transparent objects. However, reflections that occur with transparent objects can be substantially more complex. For example, light can not only hit the surface of transparent objects from the outside, but also from the inside. These internal reflections differ from the external ones in that they depend in another way on the angle of incidence and that total reflections can occur at larger angles of incidence. Like the exterior ones, internal reflections also produce mirror images of the surround. There can therefore be as many mirror images as there are reflective surfaces the light rays interact with on their way to the observer. In the case of massive transparent objects, there are typically two mirror images (see Figure 4.11a top, Figure 4.11b top). Both mirror images are superimposed additively in the image (see Figure 4.11c top). Because the light

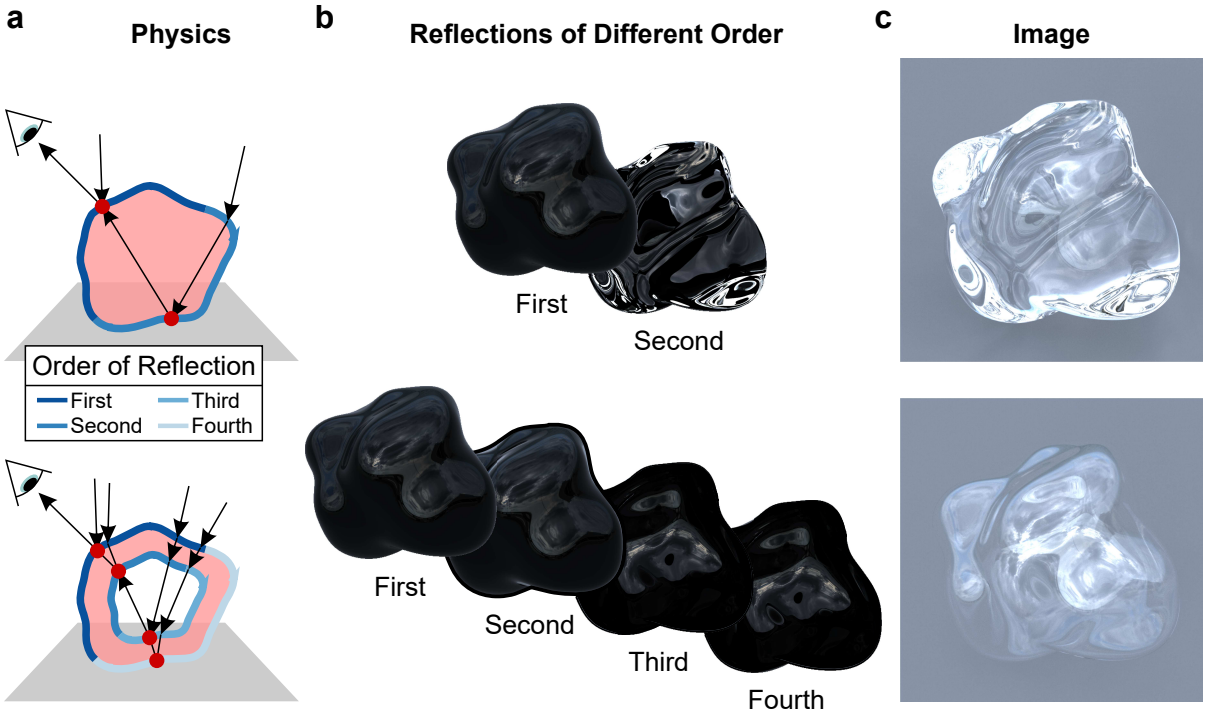

Figure 4.11: Specular reflections and mirror images of different orders caused by a massive (top) and hollow (bottom) transparent object. **(a)** With massive transparent objects (top), an observer generally sees two different reflections: One on the front surface of the object ("first-order reflection") and the other one on its rear surface ("second-order reflection"). With hollow objects (bottom), second-order reflections occur on the inner front surface. Further reflections of third and fourth order occur at the inner and outer rear surface. The different reflections are shown here schematically for one light ray each. The point at which the respective mirror image originates (i.e. where the specular reflection takes place) is highlighted by a red dot. **(b)** Example of isolated mirror images caused by reflections of different orders for a massive (top) and a hollow (bottom) transparent object (see Appendix 4.D for technical details). Note that the mirror images shown here are only rough approximations. **(c)** In the image, the mirror images of different orders are additively superimposed (see Appendix 4.A for technical details). It is therefore difficult to disentangle the different reflection components and to determine from which surface they originate.

reflected from the front surface reaches the observer directly, we denote these reflections as "first-order reflections". We call reflections at the inner back side "second-order reflections", because here the light first passes through the front surface before it reaches the observer. With hollow transparent objects, there are in addition reflections of third and fourth order (see Figure 4.11a bottom, Figure 4.11b bottom), which are also superimposed additively in the image (see Figure 4.11c bottom).

4.4.1 Transparency-specific problems in using reflections as a cue

The existence of higher-order reflections in transparent objects poses specific problems that do not occur in the case of opaque objects. We will briefly discuss the most severe of them.

First-order reflections In order to use first-order reflections as a shape cue, like in the opaque case, they first have to be isolated from higher-order reflections. In machine vision this problem has already been discussed in more detail (Morris & Kutulakos, 2007). However, such approaches cannot easily be transferred to human vision, as they typically rely on special observation conditions. In general, the isolation of first-order reflections could be facilitated by the fact that higher-order reflections are affected by the absorption properties of the object's material and therefore usually differ from first-order reflections in both intensity and spectral distribution. Furthermore, reflections of different orders move differently in the image, when the object and/or the observer moves.

Higher-order reflections Higher-order reflections must not, however, be merely regarded as a source of interference in shape recognition. In principle, they could provide further information about the shape. In particular, this could be the case with hollow objects. As long as the wall thickness is relatively small, first-order and second-order reflections are often very similar. Since both reflections are superimposed in the image, the second-order reflections can indirectly increase the visibility of first-order reflections. The similarity between the mirror images could also be used in more complex ways. For example, Shih, Krishnan, Durand, and Freeman (2015) have shown how small shifts between two mirror images could be used to detect and remove them from the image.

Higher-order reflections might also be used to estimate the shape of surface areas that are not directly visible. In principle, each reflection in the image could be used to estimate the shape of the surface on which that reflection occurs. In the case of a massive object, for example, second-order reflections could be used to estimate the shape of its rear surface. Similarly, third- and fourth-order reflections could be used to estimate the shape of the inner and outer rear surface of a hollow object. From a computational point of view, however, such a procedure would be very complex. While the light of second-order reflections is already refracted twice, the light of third- and fourth-order reflections is refracted four or six times at different surface locations of the object before it reaches the observer. Each of these refractions distorts the transmitted light pattern and changes its intensity distribution. Like in the case of (multiple) optical distortions of the background, these distortions of the mirror images interact with each other in a very complex way. In addition, light that is reflected by a specific surface might not only be refracted but also reflected several times on its previous or subsequent path. This can lead to reflections of reflections. One reason for this can be the total reflection discussed above.

Higher-order reflections depend not only on the shape of the surface that causes them, but also indirectly on the shape of all other surfaces of the object with which the light interacts. In order to gain information about the shape of a specific surface, one would have to isolate the influence of that particular surface reflection on the image from other influences. This is

difficult, because both the number of the remaining interactions and their type (reflection, total reflection, refraction) is a priori unknown.

4.5 Summary of the theoretical analyses

Our analyses revealed that the distortions of the background in the retinal image caused by transparent objects can be closely related to the shape of these objects: In simple situations there is often a direct correlation between optical magnifications and the local principal curvatures of a surface and in such cases it seems possible to estimate the shape from the direction-dependent texture densities in the image. In general, however, the relationship between background distortions, image magnifications and shape is considerably more complex. Distortions are then less related to the intrinsic curvatures of a surface, but rather to the curvatures relative to the observer. If the refraction effects are so strong that light rays cross on their way to the observer, this correlation becomes increasingly ambiguous. In cases where light is refracted at several surfaces, it is difficult if not impossible to determine the shape of one of the surfaces involved in the distortion. The extent to which background distortions contribute to the perception of shape therefore seems also to depend on the ability to identify situations in which they *do not* provide reliable information about shape.

Absorption inside a transparent material can change the intensity and spectral distribution of the transmitted light. The longer the distance the light travels inside the object, the greater the influence of absorption. In principle, the absorption-induced darkening and chromaticity changes of the background can indicate the thickness of an object and thus indirectly contribute to the recognition of the object's shape, but this darkening is often difficult to identify. Moreover, light refraction and total reflection normally weaken the correlation between darkening and thickness. Absorption could indirectly influence shape perception because it can affect the visibility of other potential shape cues.

Transparent objects can reflect light at multiple surfaces and the corresponding mirror images are superimposed in the image. The mirror images on the front surface facing the observer are identical to those of opaque objects and could therefore serve as a shape cue in a similar way. The use of the remaining mirror images would be computationally much more complex, since they are usually influenced by several interactions of the light with the surface and the material of the object.

4.6 Experiment: Testing the contribution of different cues to shape perception

In the first part of this work, we identified regularities in the image related to the shape of transparent objects. While some of the regularities known from opaque objects are no longer present (e.g. shading and texture), others remain unchanged (contour) or are present in a similar way (mirror images). Other regularities are specific to transparent objects such as background distortions due to refraction and changes in chromaticity and brightness due to absorption. The correlations between these image regularities and shape are in most situations substantially more complex than in the case of opaque objects. The different sets of potential shape cues that are available in the opaque and the transparent case suggest that the mechanisms used in shape perception depends also on these material classes. This raises two main questions: (a) How well can the shape of transparent objects be recognized and how good is the performance compared to that found with opaque objects? (b) Do the image regularities that we just analyzed theoretically play a role in the perception of shape?

To investigate these questions, we conducted an experiment in which we presented subjects with randomly shaped, blob-like objects, as they have often been used for investigating shape perception in the opaque case (e.g. Todd, 2004). To determine the perceived shape, we asked subjects to indicate the local surface orientation of the objects by adjusting small measuring probes that were projected onto their surfaces ("gauge figure task", Koenderink and van Doorn, 1992). One advantage of this method is that it allows to reconstruct the shapes perceived by the subjects. To this end, their individual settings are integrated into a global surface (Nefs, 2008; Wijntjes, 2012). In pilot experiments we also tested identification tasks that either asked the subjects to judge the equality of the shape of two objects or to identify the shape of a given object in a group of similarly shaped objects. The advantage of this kind of tasks is that they can be dealt with quickly and intuitively. However, a major disadvantage is that objects can often be identified solely on the basis of their contour. Our results in the pilot experiments indicated that shape information provided by other image regularities played only a minor role for the task. We therefore consider such methods to be inadequate for investigating differences between the perception of transparent and opaque materials. Although the unwanted influence of the contour could in principle be reduced by altering the stimulus material, for example by masking the contour or keeping it constant (so that all objects have the same outer contour), we avoided this approach, because this leads to unnatural restrictions on the type of stimuli.

Based on our computational analyses and simulation results, we used massive and hollow versions of transparent objects, since it is plausible to assume that they are processed differently. In addition, we also used objects of identical shape that were made of opaque materials, to compare the shape perception between the two material classes. Like in our computational

analyses, we limit ourselves to static stimuli and used transparent objects with smooth surfaces that had no subsurface scattering.

To test whether the image regularities that we analyzed computationally play a role in perceiving transparent objects, we manipulated their availability in the image. A possible approach would have been to present individual image regularities in isolation. However, informal observations showed that transparent objects then often lose their shape and material impression completely. In order to avoid this and to use physically plausible stimuli wherever possible, we instead omitted in each condition one image regularity from the full set.

We aimed to investigate shape perception for realistic scenes under natural viewing conditions. We therefore chose physically plausible materials (i.e. realistic refractive properties and realistic absorption spectra) and used stereoscopic stimuli. As a result, all image regularities were present in their usual form with binocular disparity. In the general discussion (see Section 4.7), we discuss the role that disparity information might play in perceiving the shape of transparent objects.

4.6.1 Stimuli

The stimuli were computer-generated stereoscopic images of randomly shaped transparent and opaque objects placed on a floor. In the transparent case, objects were either massive or hollow. The object meshes were created with the 3D modeling software Blender (Blender Foundation, 2015) and the actual stimulus images were rendered with the physically based Mitsuba renderer (Jakob, 2013). Both modeling and rendering was performed in RGB color space.

The object meshes were based on an icosahedron that was subdivided six times. The resulting icosphere consisted of 81920 triangular faces and was adjusted to a diameter of 100 mm. A total of seven deformed instances of this icosphere were created by translating its vertices along their normal direction (modifier “Displace” with parameters “Direction” = “Normal”, “Midlevel” = 0.5, and “Strength” = 1). The amount of displacement was determined by the intensity of three-dimensional Perlin noise (texture “Cloud” with parameters “Noise Basis” = “Original Perlin”, “Size” = 1, and “Depth” = 0 and options “Grayscale” and “Soft” selected). To gain different shapes, the noise was probed at different locations. We avoided locations that would have led to shapes with extensive self-occlusions, because this would have unnecessarily complicated the later reconstruction of perceived surface shapes. The seven meshes used in the experiment are shown in Figure 4.12. While the position of the objects in the scene remained constant, the vertical position of the floor was adjusted to compensate for the varying vertical extent of the objects. Hollow objects had a wall thickness of 1 mm and were created by eroding their interior without changing the outer shape (modifier “Solidify” with parameters “Thickness” = 2 and “Offset” = −1 and options “Even Thickness”

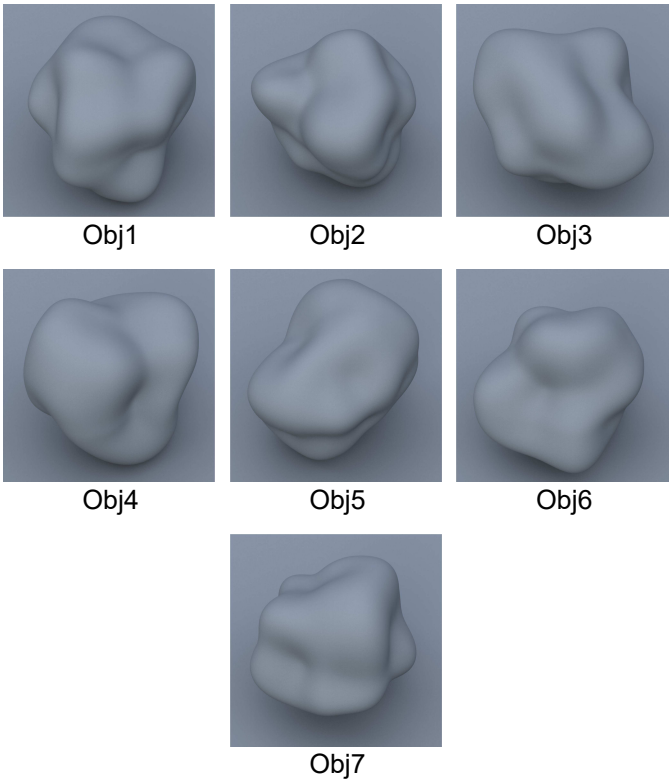

Figure 4.12: The seven blob-like object meshes used in the experiment. The meshes were designed to resemble the ones that were used in previous work on shape perception.

and "High Quality Normals" selected). After modeling, the objects were exported and the scene rendered with the Mitsuba renderer. In the rendering, all objects were surrounded by (and, if hollow, filled with) non-absorptive air (refractive index $R \approx 1$).

In eight different cue conditions, four for opaque objects and four for both massive and hollow transparent objects, we manipulated the availability of different known or potential shape cues (see Figure 4.13). For all except one cue condition, this was achieved by choosing appropriate material properties for both the object and its background. In the remaining cue condition ("Mirr−"), we explicitly manipulated the image generation. For both transparent and opaque objects, we defined base conditions that contained all cues that were manipulated in the remaining conditions.

In the base condition for the massive transparent case ("Trns:Mass:Full"), the objects were made of red tinted acrylic glass with a smooth surface (Mitsuba plugin "dielectric", refractive index $R = 1.49$, absorption coefficient $a_{\text{Mass}} = \text{RGB}(.0048, .0072, .0072)\,1/\text{mm}$). In the base condition for hollow transparent objects ("Trns:Holl:Full"), a higher absorption coefficient was used ($a_{\text{Holl}} = \text{RGB}(.176, .264, .264)\,1/\text{mm}$) in order for massive and hollow objects to appear similarly tinted. In the remaining three cue conditions for the transparent case, potential shape information from distortions, darkening, and mirror images were omitted individually. To omit optical distortions of the background (cue condition "Dist−"), we changed the color

of the background to a uniform gray with a reflectance of RGB(.2, .2, .2). To omit darkening due to absorption (cue condition "Dark−"), the object's material was set to not absorb any light ($a_{\text{Dark}-}$ = RGB(0, 0, 0) 1/mm). To omit mirror images (cue condition "Mirr−"), specular reflections at any of the object's surfaces were disabled during the rendering process.

In the opaque base condition ("Opq:Full"), the objects were made of plastic (polypropylene) with a smooth surface (Mitsuba plugin "plastic", refractive index $R = 1.49$) and had a red granite-like texture based on three-dimensional Perlin noise (texture "Cloud" with "Noise Basis" = "Original Perlin", "Size" = 1.5, "Depth" = 2 and options "Grayscale" and "Soft" selected). The reflectance of the texture ranged from RGB(.144, .096, .096) to RGB(.336, .224, .224). Each object was textured individually, so that the intrinsic texture density remained constant despite the different shapes. In two additional cue conditions for the opaque case, shape information from texture and mirror images were omitted individually. In the fourth condition, information from mirror images were presented in isolation from the other cues. This condition was implemented because it corresponds to the large material class of reflective metals that has already been studied in more detail in some of the works on the perception of the shape of opaque objects. To omit texture information (cue condition "Tex−"), the objects' surface color was changed to a uniform red with a reflectance of RGB(.24, .16, .16). To omit mirror images (cue condition "Mirr−"), the objects were made of an ideal diffuse material with a Lambertian reflectance (Mitsuba plugin "diffuse"). To isolate mirror images (cue condition "Mirr+"), the objects were made of a metal-like material that reflects incoming light specularly (Mitsuba plugin "conductor", reflection coefficient r = RGB(.72, .48, .48)).

Unless otherwise stated, the floor below the objects showed a gray graph paper that was made of two superimposed grid textures of different size. This texture was well suited to depict a wide range of optical magnifications and compressions as they were caused by the different refractive and reflective materials used here. The background was made of an ideal diffuse material with a Lambertian reflectance (Mitsuba plugin "diffuse"). The width of the two grids were 1 mm and 10 mm, respectively. The reflectance of the grid lines (RGB(.28, .28, .28)) was slightly higher than the one of their surround (RGB(.12, .12, .12)).

The scene was illuminated by an environment map (Mitsuba plugin "envmap"). The illumination texture was a high dynamic range image (color depth 16 bit/channel) of a natural daylight outdoor scene with a partly cloudy sky (Yimm & Bell, 2008). This environment map was considered a typical representative of a realistic and natural ambient lighting.

The camera settings were chosen to correspond to the actual experimental setup (distance to the center of the object 400 mm, vertical field of view 44.10°, lateral offset ±32 mm). Thus, the stimuli appeared in virtually the same way as a corresponding real scene.

The stimuli were first rendered as high dynamic range images (color depth 16 bit/channel) with the Mitsuba renderer (extended volumetric path tracer with "maximum path depth" = 64; Hammersley sampler with 2048 samples/px; Gaussian reconstruction filter with "standard

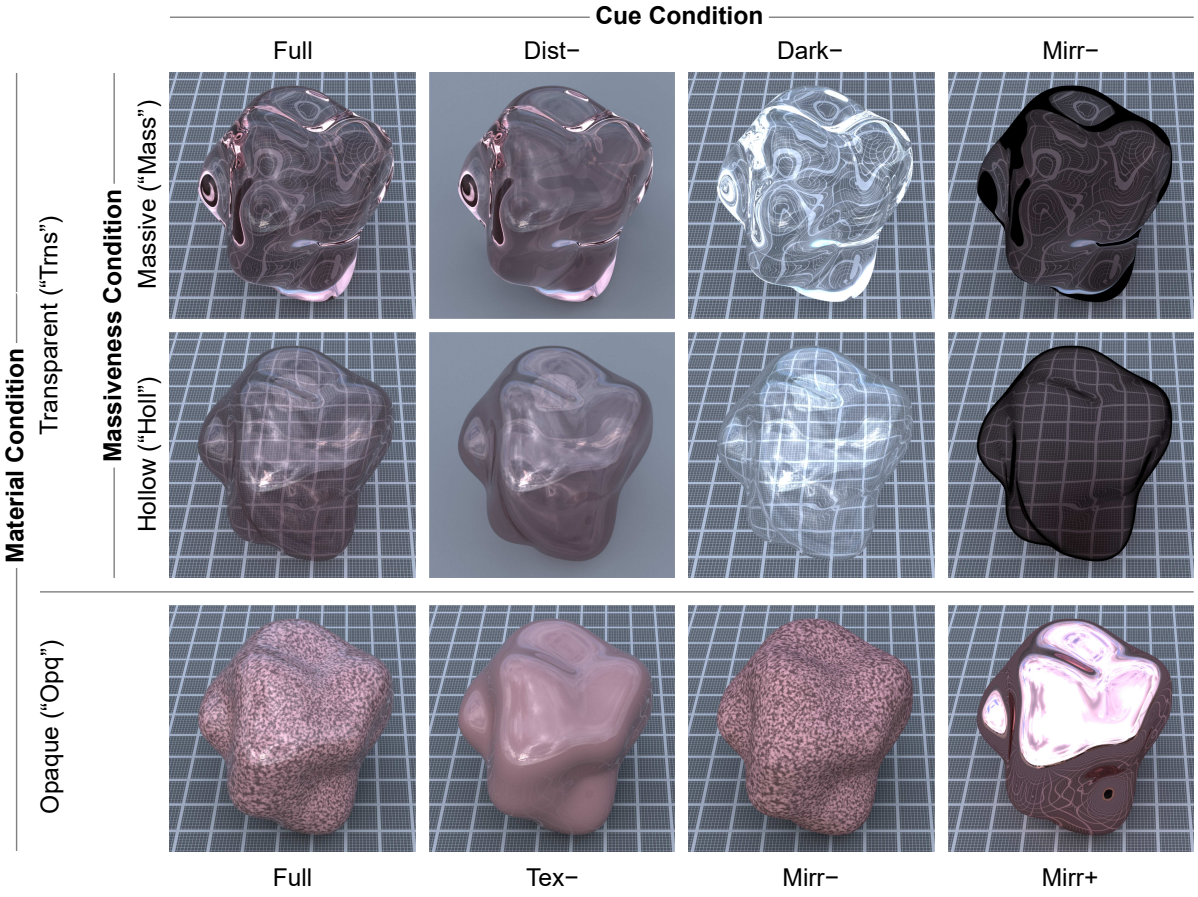

Figure 4.13: Stimulus conditions used in the experiment. The material of the objects was either transparent (top two rows, "Trns") or opaque (bottom row, "Opq"). The transparent objects were either massive (top row, "Mass") or hollow (second row, "Holl"). Based on three base conditions (leftmost column, "Full") one potential cue was omitted in each of the remaining cue conditions. For the transparent objects, this was either background distortions ("Dist−"), darkening from absorption ("Dark−") or mirror images ("Mirr−"). For the opaque objects, this was either texture ("Tex−") or mirror images ("Mirr−"). In addition, a metal-like opaque object was presented in which the mirror images were isolated ("Mirr+"). The name of each stimulus condition is given by its abbreviated material, its massiveness (if applicable), and the respective cue condition (e.g. "Trns:Holl:Dist−" for the stimulus condition that shows a hollow transparent object without background distortions). Note that here, only the stimulus images intended for the right eye are shown and that they are trimmed for presentation purposes.

deviation" = 0.5; image size = 839 × 1200 px). Note that more complex effects resulting from the dispersion of light or its polarization were not taken into account.

To compensate for the limited dynamic range of the display device used in the experiment, the stimuli were tonemapped to low dynamic range images (color depth 8 bit/channel) according to the procedure described by Reinhard and Devlin (2005). All stimuli were tonemapped with the same set of parameters. This refers to both the initial parameters (contrast = 0.1, intensity = 1.5, chromatic adaptation = 0.1, light adaptation = 1) and the implicit image-dependent parameters that were first gained for one high dynamic range stimulus that contained high luminance values (drawn from the opaque cue condition "Mirr+") and subsequently used for the remaining stimuli. Due to the limited horizontal field of view of the mirror stereoscope used in the experiment, the stimulus images for the right and left eye were trimmed at their right and left edges, respectively. The final size of a half image was 720 × 1200 px. To slightly increase the contrast of the images, they were gamma corrected with an exponent of $\gamma = 1.2$, which is slightly lower than the gamma value (1.6) that was used by Reinhard and Devlin (2005).

4.6.2 Subjects

A total of 42 subjects, 38 of them female, participated in the experiment. Their age ranged from 18 to 34 years. All subjects were naive as to the purpose of the experiment. They reported normal or corrected-to-normal visual acuity, and showed no color vision deficiency, as tested by Ishihara plates (Ishihara, 1969).

4.6.3 Procedure

To keep the duration of the experiment within reasonable bounds, each of the 42 subjects performed only a subset of all conditions. Each subject was assigned to one of two groups and each group was presented with seven out of 12 stimulus conditions consisting of massive and hollow transparent and opaque objects in different cue conditions. The first group was presented with three opaque ("Opq:Full", "Opq:Tex−", and "Opq:Mirr−") and all four massive transparent stimulus conditions ("Trns:Mass:Full", "Trns:Mass:Dist−", "Trns:Mass:Dark−", and "Trns:Mass:Mirr−"). The second group was presented with two opaque ("Opq:Full" and "Opq:Mirr+"), one massive transparent ("Trns:Mass:Full") and four hollow transparent stimulus conditions ("Trns:Holl:Full", "Trns:Holl:Dist−", "Trns:Holl:Dark−", and "Trns:Holl:Mirr−"). The base cue conditions for the opaque and the massive transparent case were shared by both groups to control for any systematic effects between the groups. In both groups, the seven randomly shaped objects were balanced across the seven stimulus conditions and the 21 subjects. As a result, every object was combined with every stimulus condition. Although each subject was presented with all seven objects, each object was presented in only one

stimulus condition. This ensured that subjects did not see objects of identical shapes in different stimulus conditions.

The stimuli were presented on an LCD screen (Eizo ColorEdge CG243W, Eizo Corporation, Hakusan, Japan; display area 518.4 × 324.0 mm; resolution 1920 × 1200 px with 3.704 px/mm; color depth 8 bit/channel) and were viewed through a mirror stereoscope (SA200 Screen-Scope Pro, Stereo Aids, Albany, Australia; optical viewing distance = 400 mm; interocular distance = 64 mm). The size of each half image (720 × 1200 px) corresponded to 194.4 × 324 mm on the screen.

In each trial, the subjects were asked to indicate the orientation of the normal ("gauge figure task", Koenderink and van Doorn, 1992) at one of 160 surface points (see Figure 4.14). Inputs were made by mouse and keyboard. The measurement points were evenly distributed in a triangular grid so that they fitted into the respective object area. Because the outlines of the various objects differed, the resulting number of triangular faces varied slightly between 272 and 275. At the rim of the objects, the inclination of the surface to the observer is maximal (i.e. its slant is 90°). To avoid trivial settings, the measurement points were located at least 5 px away from the rim. The gauge figure was highlighted in green (RGB(0, .98, .60)), had a maximum base diameter of 24 px, and a maximum rod length of 12 px, with a line width of 2 px.

A problem when using the gauge figure method in stereoscopic stimuli is to decide at what depth the gauge figure should be positioned. If it is positioned at an arbitrary depth (e.g. at the image plane of the screen), this may make the adjustment more difficult, as the gauge figure may not necessarily appear to lie on the surface of the object, but in front or behind of it. If it is positioned correctly on the actual surface of the object, its perceived depth may indirectly provide information about the shape of the object and thus interfere with the information provided by other cues. In order not to provide the subjects with stereoscopic cues to the depth of the gauge figure, we therefore presented it to the right eye only. In a preliminary experiment, in which we used this monocular presentation mode we did not observe the systematic overestimation of the perceived surface slant that was found by Bernhard, Waldner, Plank, Solteszova, and Viola (2016), who combined a gauge figure without disparity with stereoscopic stimuli.

To gain experience with the gauge figure task, subjects performed numerous practice trials, until they felt up to the task. The shapes of the objects shown in the practice phase differed from those used in the experiment. In the actual experiment, each subject performed 1120 trials in a randomized order (7 stimulus conditions × 160 measurement points). This resulted in three repetitions (each by a different subject) of the 160 individual measurements belonging to each of the 84 combinations of seven object meshes and 12 stimulus conditions. In the 14 combinations involving the two base conditions "Opq:Full" and "Trns:Mass:Full", three additional repetitions were made. On average, a subject required roughly 65 minutes

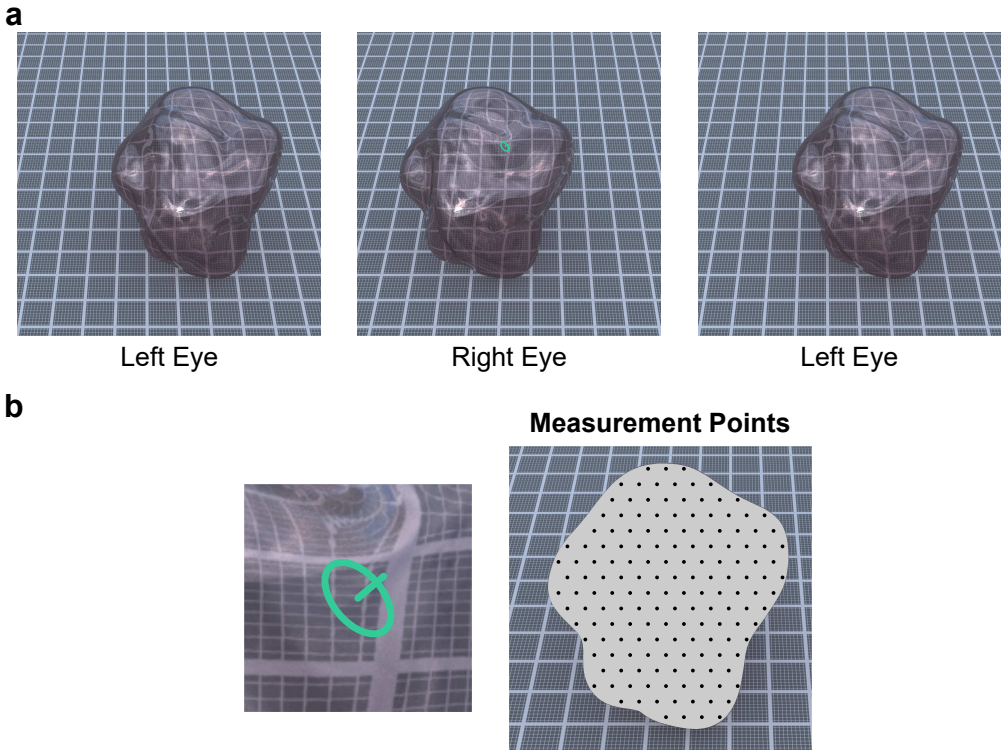

Figure 4.14: Stimulus example and measurement points. **(a)** Example of stereoscopic stimulus images showing a hollow transparent object "Obj1" in its base condition (stimulus condition "Trns:Holl:Full"). The gauge figure was presented to the right eye only and remained visible throughout the adjustments. The images shown here are meant for crossed fusion (right image pair) or parallel fusion (left image pair). In the experiment, the perspective properties used in the rendering of the stimuli and the geometry of the mirror stereoscope were compatible (this included the viewing distance, the field of view and the lateral stereo offset). Note that the brightness and contrast of the images shown here have been increased. Furthermore, the images were cropped vertically. **(b)** The left side shows a closeup of the gauge figure shown in Figure 4.14a. The right side indicates the 160 measurement points at which the gauge figure was presented in different trials using the example of object "Obj1".

for performing all of the trials. Due to the additional practice phase and possible rests, the experiment was divided into multiple sessions depending on the speed of the subject.

After the experiment, the subjects were asked about their material and massiveness impression. To this end, they were presented with printed copies of stimulus images that showed an object similar to that used in the experiment under different stimulus conditions. The subjects were asked to indicate whether the respective object material appeared transparent or opaque to them. In addition, the subjects of the second group were asked whether the object shown in the respective conditions appeared solid or hollow.

4.6.4 Results

The subjects' settings were analyzed in different ways that each focused on a specific aspect of the perceived shape and its deviation from the actual shape. We started by considering each

setting as an independent measurement of local surface orientation. To evaluate the relative performance in the 12 stimulus conditions, we compared means across different local error measures observed under these conditions are compared. In a second step, we reconstructed the surfaces shapes perceived by the subjects from their local settings. Here the error is given by the deviation of the reconstructed surface from the actual one. This more global approach allows to distinguish between qualitative and quantitative errors and to analyze systematic misjudgments of the local shape.

Since we found no systematic differences between the two base conditions "Trns:Mass:Full" and "Opq:Full" that were shared between the two groups of subjects, the redundant settings of the second group were discarded in order to maintain equal group sizes.

4.6.4.1 Analysis of Normal Errors

An obvious way to evaluate the local gauge figure settings is to compute the "normal error", i.e. to directly compare the unit normal vectors indicated by the gauge figures with the veridical unit normal vectors (see Figure 4.15a). Figure 4.15b shows that this normal error Δn is considerably higher for the massive and hollow transparent base conditions ($\overline{\Delta n}_{\text{Trns:Mass:Full}} = 38.74°$ and $\overline{\Delta n}_{\text{Trns:Holl:Full}} = 28.39°$, respectively) than for the opaque base condition ($\overline{\Delta n}_{\text{Opq:Full}} = 14.11°$). In the transparent case, the omission of background distortions and mirroring had opposite effects for massive and hollow objects. Without background distortions (cue condition "Dist−"), the normal error decreases by 3.08° for massive, but slightly increases by 1.10° for hollow objects. In contrast, the normal error increases by 8.96° if mirroring is omitted for massive objects, while it decreases by 3.15° for hollow ones ("Mirr−"). The omission of absorption-induced darkening ("Dark−") has negative effects for both the massive and the hollow case (2.19° and 0.56°, respectively). Although the normal error is lowest in the opaque base condition, it increases to an extent that is similar to that found for hollow transparent objects if information from texture is missing ("Tex−") or the object is made of a fully reflective material ("Mirr+"). In contrast to the transparent case, the omission of mirroring had no effect on the normal error in the opaque case ("Mirr−").

4.6.4.2 Systematic and random local errors

A more sophisticated approach to evaluate the local gauge figure settings is to analyze the variance of the adjusted normals about the veridical ones and decompose it into accuracy and precision components, to distinguish between systematic and random errors (see Figure 4.16a). We calculated the total variance and this decomposition separately for each point of measurement, each object and each stimulus condition. Figure 4.16b shows the pattern of the total normal variances across the different stimulus conditions, which closely resembles the pattern found for the normal error Δn (cf. Figure 4.15b). The variances are considerably

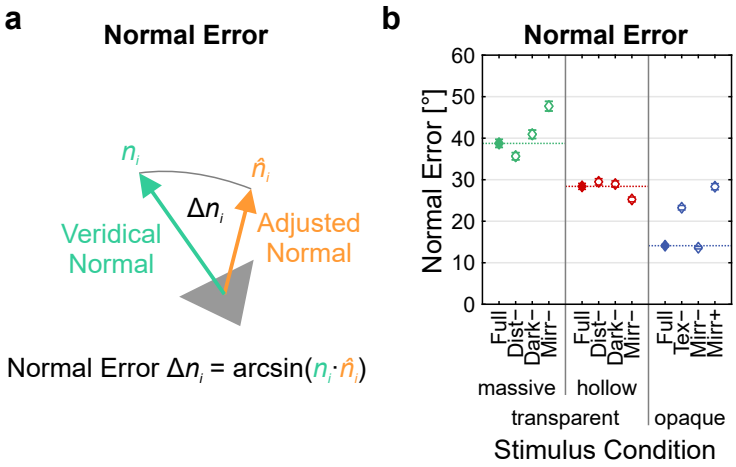

Figure 4.15: Analysis of the angular normal error. **(a)** The normal error Δn_i is defined as the (unsigned) angular difference between the unit normal vector indicated by a gauge figure setting i ($\hat{n}_i$, yellow) and the veridical unit normal vector (n_i, green). Note that both n and $\hat{n}$ are defined with respect to the image plane. Thus, n is not identical to the veridical surface normal in the world space. **(b)** Angular normal error (±95 %-CI) for each stimulus condition, averaged across all objects, points of measurement and subjects. The error levels of the base conditions ("Trns:Mass:Full", "Trns:Holl:Full", and "Opq:Full") are emphasized by dashed horizontal lines. Due to the restricted adjustment range of the gauge figure, the maximum averaged normal error $\overline{\Delta n}_{\max}$ that would occur if the subjects make the greatest possible error for each of their setting depends on the distribution of the veridical surface normals. For the stimuli used in this experiment, $\overline{\Delta n}_{\max} = 112.52°$. If the subjects' settings were random, the expected averaged normal error would be $\overline{\Delta n}_{\text{random}} = 90°$.

higher for massive transparent objects than for hollow ones. The lowest variances can be found for opaque objects. In addition, the data indicate that in the transparent conditions, the accuracy variance tends to take a greater share of the total variance than the precision variance (roughly about 60 %). Apart from the case of stimuli without texture, the relative contribution of accuracy and precision variance tends to be more balanced for opaque objects throughout (see Figure 4.16c, which explicitly shows the share of the accuracy variance of the total variance for different stimulus conditions).

The relative influence that each potential cue had on the settings is shown in Figure 4.17. It turns out that the negative influence of optical background distortions that we found for massive transparent objects is mainly due to an increase of the systematic error (i.e. the accuracy variance). In contrast, its slightly positive influence for hollow objects is caused by a rather equal decrease in both systematic and unsystematic variance. For both massive and hollow transparent objects, mirroring affected the accuracy variance to a larger extent than the precision variance. However, the influence was more positive in the massive transparent case than it was negative in the hollow case. Its positive effect was even stronger than the positive effect of the texture cue in the opaque case. In contrast to the transparent case, the mirroring cue had virtually no influence on shape perception in the opaque case. While the positive effect of absorption-induced darkening observed for massive transparent objects is due to a decrease of both systematic and unsystematic errors, the small effect for hollow objects is due to a decrease of the accuracy variance only.

4.6.4.3 Effect of contour information

Besides the shape cues that were manipulated in the experiment, the objects' contour is always present as an additional shape cue. To examine how strong the relative influence of the contour is, we plotted the size of the angular normal error Δn against the distance of the measurement points from the contour. Figure 4.18 shows that in the case of opaque objects, the normal error is hardly affected by the proximity of the measurement points to the object contour. For transparent objects, however, the normal error tends to increase with contour distance. This trend can be found for both massive and hollow transparent objects, whereby the difference in the error level is approximately constant for all contour distances. For very small contour distances, the error level found for hollow transparent objects almost decreases to the level found for opaque objects.

4.6.4.4 Local slant/tilt errors

Previous work has shown that subjects tend to underestimate the surface slant, i.e. the angle between the perceived surface normal and the line of sight (e.g. Bernhard et al., 2016; De Haan, Erens, & Noest, 1995; Koenderink & van Doorn, 1992; Todd, Oomes, Koenderink, & Kappers,

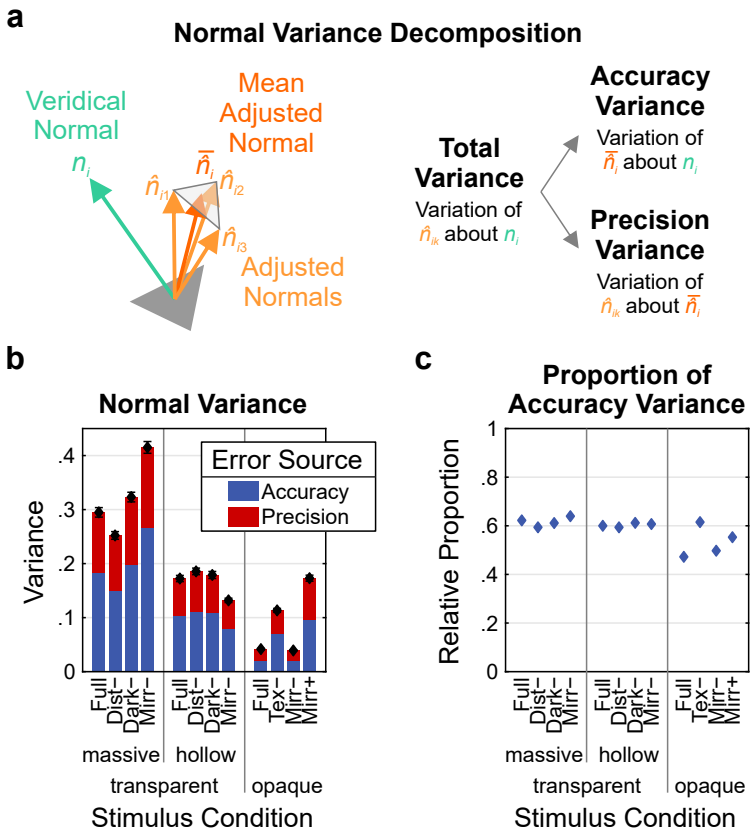

Figure 4.16: Analysis of systematic and random variance of the normal. Since normals are directions, the decomposition was based on spherical variance measures (see Appendix 4.E for details). **(a)** The total variance of the normals adjusted for a particular point of measurement, object and stimulus condition, was decomposed into accuracy and precision components to distinguish systematic from random errors. The precision variance describes the variation of the k individual settings ($\hat{n}_{ik}$) made by three subjects about their mean $\overline{\hat{n}}_i$, where i denotes a specific measurement. The accuracy variance describes the variation of the mean setting $\overline{\hat{n}}_i$ about the corresponding veridical normal n_i. To compare different cue conditions we pooled the variances across all points of measurements and objects used in the experiment. **(b)** Accuracy and precision components of the total variance (±95 %-CI). The value of the total variance can be between 0 and 1, where 1 means that the adjusted normals are equally distributed in all directions. **(c)** Relative proportion of the accuracy variance in the total variance for each stimulus condition.

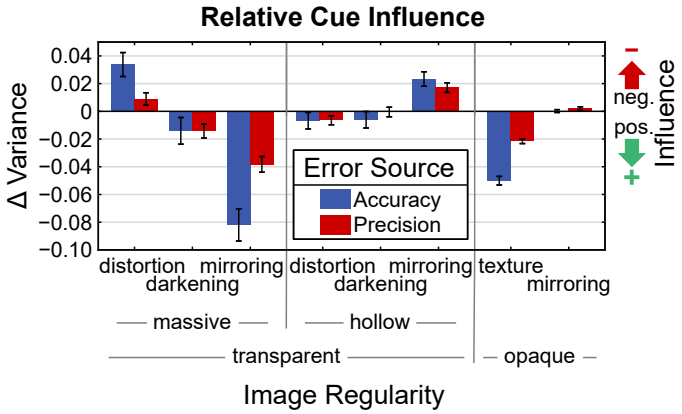

Figure 4.17: Deviation of accuracy and precision variance (±95 %-CI) in cue conditions with omitted cues from the values in their respective base condition. Positive values indicate that the existence of the respective image information increases the variance, which means that it has a negative influence on shape perception. Note that just because a potential cue has no influence on the normal variance, this does not necessarily mean that it is irrelevant for shape perception (see discussion, Section 4.6.5).

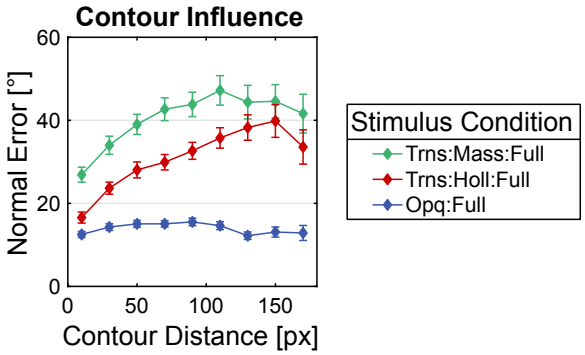

Figure 4.18: Angular normal error (±95 %-CI) as a function of the distance between the respective measuring point and the contour of the object, shown for the transparent and opaque base conditions. The displayed values correspond to an interval of ±10 px and are averaged across all objects, points of measurement and subjects. A contour distance of 185 px roughly corresponds to the average radius of the objects in the image.

2004). To test for this effect and further potential differences between the stimulus conditions, we reparametrized the normals indicated by the gauge figure in terms of spherical slant and tilt coordinates (see Figure 4.19a). Stevens (1983) argued that this parametrization corresponds well to how the visual system represents the orientation of surfaces.

The general pattern of slant and tilt errors (see Figure 4.19b) is similar to that found for the angular normal error Δn (cf. Figure 4.15b). However, the tilt error tends to be substantially larger than the slant error. The interpretation of surface tilt is difficult, if adjusted and veridical surface normal are both close to the line of sight, because then even small angular differences can lead to large tilt errors. However, for the convex objects used in the experiment such cases are very rare and thus cannot explain the relatively large tilt errors. As expected, the slant bias was negative for every stimulus condition, which means that on average, the slant was underestimated for both transparent and opaque objects (see Figure 4.19c). In addition, the degree of this underestimation differs between the stimulus conditions. The corresponding pattern of results is roughly inversely proportional to the previous error measures. On average, the underestimation of slant was most pronounced for massive transparent objects and smallest for opaque objects. The general underestimation of the slant contradicts the results of Bernhard et al. (2016), who found a systematic overestimation of the perceived surface slant when a monoscopic gauge figure is presented with stereoscopic stimuli. In the tilt dimension, no systematic bias was found (i.e. $\mathrm{B}\tau \approx 0$ for each stimulus condition).

4.6.4.5 Reconstruction of perceived surfaces

One drawback of local error measures like the ones discussed so far is that they do not directly indicate what surface shape subjects perceived when they were presented with a specific object. In particular, it is possible that surfaces that are perceived differently lead to identical averages of local errors. Consider, for example, a flat surface that is erroneously perceived as concave by one subject and as convex by another. Obviously, the average normal error of both subjects can nonetheless be the same. To allow for a more global interpretation of the subjects' settings, we integrated the local gauge figure data into triangular meshes that were meant to reflect the perceived object shapes (see Figure 4.20a). This surface reconstruction was first proposed by Koenderink and van Doorn (1992). The actual procedure is described by Nefs (2008) and Wijntjes (2012). Because the reconstructed surfaces are based on the adjusted gauge figures, we will refer to them as "adjusted surfaces". Since surface reconstruction is performed in the image space, we subsequently transformed the reconstructed surfaces into the world space, to compare them with the veridical ones (see Figure 4.20b and Appendix 4.F).

Figure 4.19: Analysis of the normal error with respect to the line of sight. **(a)** An alternative way of analyzing the normal error is to take the viewing direction of the observer into account, by parameterizing both adjusted and veridical normals in terms of spherical slant and tilt. The slant component σ is the angle between the normal and the line of sight ($\sigma \in [0°, 90°]$). The tilt component τ describes the orientation of the normal in the image plane ($\tau \in]-180°, 180°]$). Accordingly, the deviation between adjusted and veridical normals can be decomposed into the slant error $\Delta\sigma_i = |\hat{\sigma}_i - \sigma_i|$ (blue) and the tilt error $\Delta\tau_i = |\hat{\tau}_i - \tau_i|$ (red), where i denotes a specific measurement. Systematic over- or underestimations of the two parameters are given by the slant bias $B\sigma_i = \hat{\sigma}_i - \sigma_i$ and the tilt bias $B\tau_i = \hat{\tau}_i - \tau_i$. **(b)** Slant error $\Delta\sigma$ (left) and tilt error $\Delta\tau$ (right) for each stimulus condition, averaged across all objects, points of measurements and subjects (±95 %-CI). **(c)** Slant bias $B\sigma$ (±95 %-CI) for each stimulus condition, averaged across all objects, points of measurements and subjects.

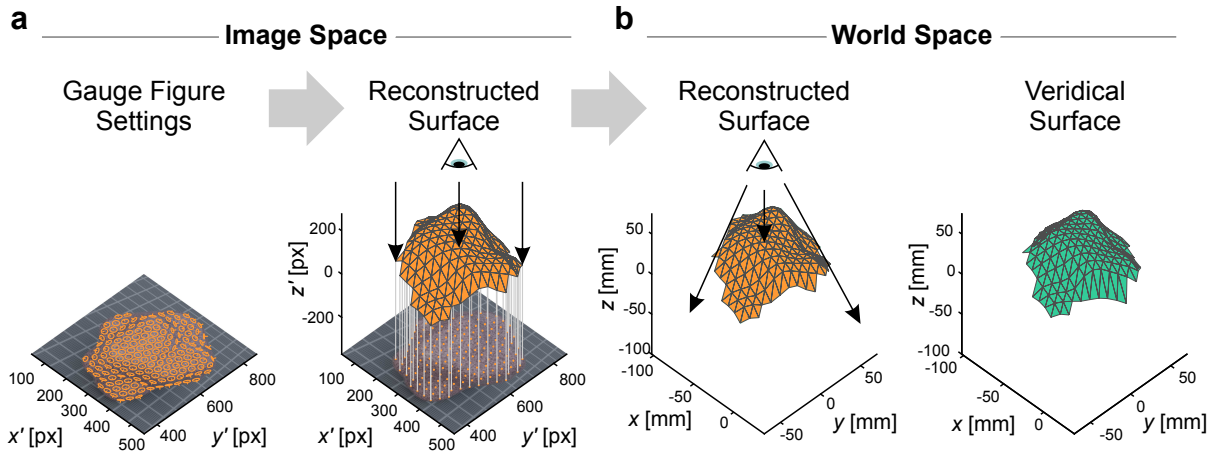

Figure 4.20: Exemplary reconstruction of the perceived surface and comparison with the corresponding veridical surface for the Object "Obj1" in the stimulus condition "Opq:Full" for one subject. **(a)** To analyze the surface shapes perceived by the subjects, their individual gauge figure settings (left diagram) were integrated to triangular meshes (Koenderink & van Doorn, 1992; Nefs, 2008; Wijntjes, 2012). Basically, this surface reconstruction involves adding a third dimension to the image space and assigning to each point of measurement a depth value that fits the data best (right diagram). Because extreme gauge figure settings with a slant value of 90° can lead to reconstructed surfaces with infinite depth expansion, we limited the range of the adjusted slant values so that $\hat{\sigma}_i = \min(\hat{\sigma}_i, 89°)$. Note that the reconstructed depth values are defined along the respective viewing directions of the surface points (black arrows). While different viewing directions run parallel to the Z' axis in the image space, they diverge in world space due to the perspective projection. **(b)** To compare the reconstructed surfaces with the veridical ones (right diagram), we subsequently transformed them into the world space (left diagram; see Appendix 4.F for details). To this end, the reconstructed surfaces were anchored at a specific distance to the observer, assuming that their centers of gravity coincide with the respective veridical surfaces. This corresponds to the assumption that the subjects were able to accurately judge the overall distance of the objects. For the analysis of the data, the resolution and range of the veridical mesh was reduced to match that of the reconstructed surface.

4.6.4.6 Analysis of Surface Similarity

To compare adjusted and veridical surfaces, we first analyzed the depth differences between them, which we call Δd (see Figure 4.21a). In contrast to the local error measures discussed previously, the largest depth errors do not occur for massive but for hollow transparent objects (see Figure 4.21b, black points). Furthermore, the variances found in these conditions are substantially higher than for most of the remaining conditions. To determine whether this is due to the susceptibility of the reconstruction algorithm to extreme individual gauge figure settings, we also analyzed the depth error for surfaces that were reconstructed from the averaged adjusted normals $\overline{\hat{n}}_i$ instead of the individual adjusted normals $\hat{n}_{ik}$. The underlying idea is that the reconstruction algorithm delivers more reliable results if the reliability of its input increases. In fact, the depth errors of the surfaces reconstructed in this way are smaller in every stimulus condition. Furthermore, the depth error for hollow transparent objects is now smaller than that for massive ones (see Figure 4.21b, blue points). This contrasts to the results just discussed, but is consistent with the result pattern we found for the normal error Δn (cf. Figure 4.15b). Another indication that the deviating result pattern of individually reconstructed surfaces is mainly due to the susceptibility of the reconstruction to outliers is that the depth deviation between the differently reconstructed surfaces is most pronounced for hollow transparent objects. While one could assume that this is due to the fact that the precision of the subjects is particularly low in these conditions, the variance decomposition of the normal error showed that this is not the case. Instead, the precision variance for hollow transparent objects was lower (i.e. the precision was higher) than for massive transparent objects. This shows that an analysis based on individually reconstructed surfaces runs the risk of overestimating the actual depth errors. There exist other measures for the difference between two surfaces, such as the modified Hausdorff distance (Dubuisson & Jain, 1994), that presumably are less susceptible to outliers. However, an analysis of this measure showed the same pattern of results that we found for the depth error Δd.

4.6.4.7 Qualitative and quantitative shape errors

There are cases where measures that refer to differences between adjusted and veridical surface points do not adequately represent the goodness of shape perception. If, for example, convex and concave surface patches are correctly perceived as such, but their strength of curvature is under- or overestimated, adjusted and veridical surfaces can differ, although the type of local shape is judged correctly. Furthermore, as a consequence of a global surface reconstruction, large normal errors at one point can indirectly lead to large depth errors at points where the normal errors are actually lower. It therefore seems more sensible to compare adjusted and veridical surfaces in terms of their qualitative and quantitative shape similarity. To this end, we analyzed local shape indices and curvedness values, two measures

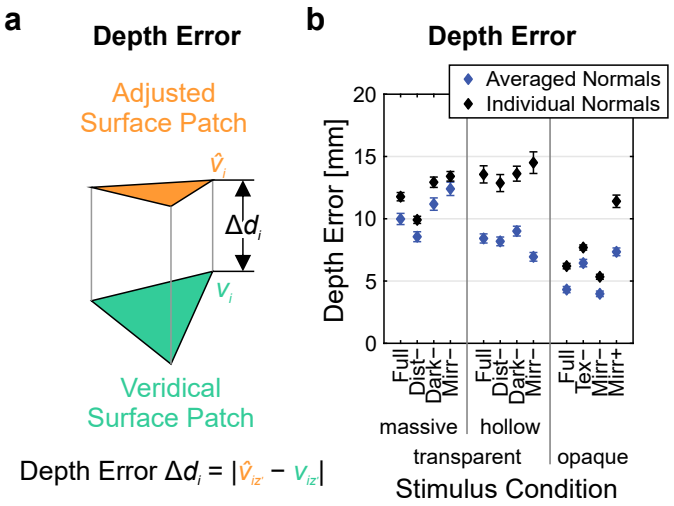

Figure 4.21: Analysis of the depth error between adjusted and veridical surfaces. **(a)** We defined the depth error Δd_i as the absolute distance between an adjusted vertex $\hat{v}_i$ of the reconstructed surface and the corresponding vertex v_i of the veridical surface, where i denotes a specific vertex of the reconstructed surface. Since the distance is measured along the respective viewing direction on which both vertices lie, Δd_i corresponds to the difference of the z' coordinates of the corresponding vertices in the image space (cf. Figure 4.20). **(b)** Depth error (±95 %-CI) for each stimulus condition, averaged across all objects, points of measurements and subjects (black data points). The blue data points correspond to the depth error (±95 %-CI) for surfaces that were reconstructed from the mean normals $\bar{n}$ instead of the individual normals $\hat{n}$, averaged across all objects and points of measurements.

that describe the type of local surface shape and the strength of local curvature, respectively (Koenderink & van Doorn, 1992). The calculation of the principal curvatures was done using a procedure proposed by Rusinkiewicz (2004).

One way to analyze the accuracy of the reconstructed shapes is to correlate adjusted and veridical shape index and curvedness values. The top row of Figure 4.22a shows the correlations between adjusted and veridical shape indices for the transparent and opaque base conditions. The correlation is highest for the opaque base condition ($R = .75$, right), considerably weaker for the hollow transparent ($R = .32$, center) and almost absent in the massive transparent ($R = .13$, left) base condition. If the correlation coefficients of all stimulus conditions are considered, it turns out that the resulting pattern is roughly inverted relative to the pattern of the normal error Δn (see Figure 4.22b left). A similar pattern can be found for the curvedness (see Figure 4.22a bottom, Figure 4.22b right). The correlation is strongest for objects of the opaque base condition ($R = .57$) and substantially lower for the hollow and massive transparent base conditions ($R = .21$ and $R = .13$, respectively).

Another way to analyze the accuracy of the reconstructed shapes is to consider the error distribution of the shape index and curvedness values. We define the shape index error as the absolute difference between the adjusted and the veridical shape indices ($\Delta s_i = |\hat{s}_i - s_i|$, where i denotes a specific vertex of the reconstructed surface). The larger Δs, the more the local

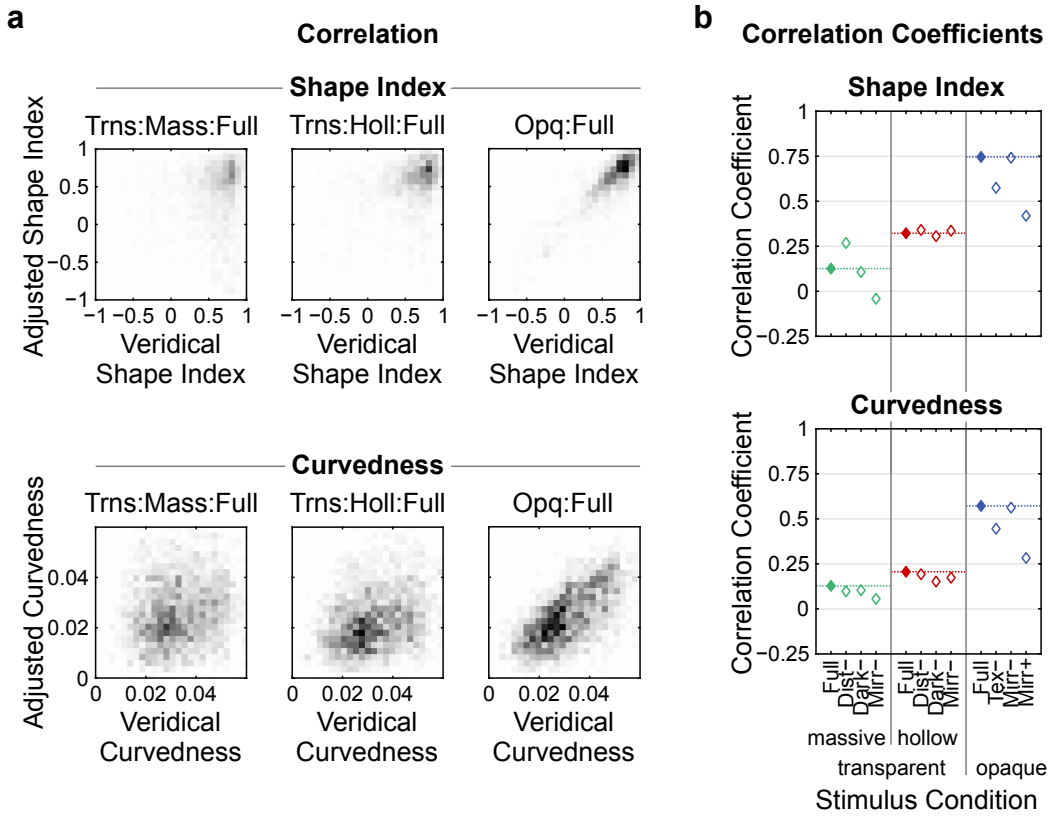

Figure 4.22: Analysis of the correlation between adjusted and veridical shape indices and curvedness values. **(a)** Bivariate histogram of adjusted (ordinate) and veridical (abscissa) shape indices ($\hat{s}$ and s respectively, top row) and curvedness values ($\hat{c}$ and c respectively, bottom row) for all transparent and opaque base conditions (columns), pooled across all objects and points of measurements. As negative shape indices are less common for the overall convex objects used in this experiment, most of the data points accumulate at positive shape index values. **(b)** Correlation coefficients R for the correlation between adjusted and veridical shape indices (left) and curvedness values (right) for all stimulus conditions, pooled across all objects and points of measurements.

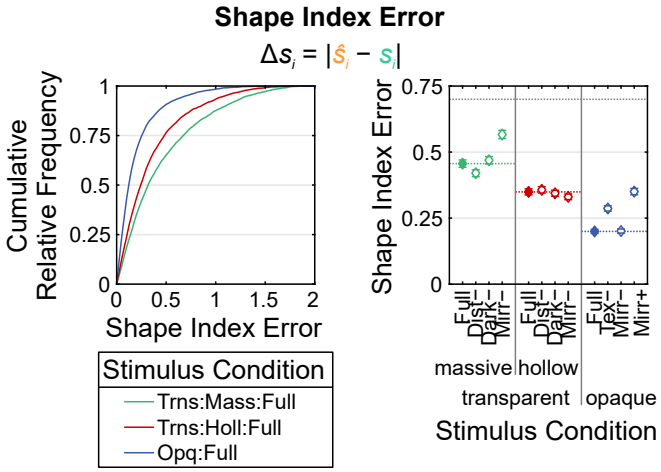

Figure 4.23: Analysis of the shape index error $\Delta s_i = |\hat{s}_i - s_i|$, where $\hat{s}$ denotes the local shape index of the adjusted, s the local shape index of the veridical surface and i a specific vertex of the reconstructed surface. The left diagram shows the cumulative frequency distribution of Δs for the transparent and opaque base conditions, pooled across all objects, points of measurements and subjects. The right diagram shows Δs (±95 %-CI) for all stimulus conditions, averaged across all objects, points of measurements and subjects. Note that due to the restricted range of the shape index, the maximum $\Delta s_{\max} = 2$ can only occur for locations where the veridical shape index is either -1 or 1. The maximum averaged shape index error $\overline{\Delta s}_{\max}$ therefore depends on the distribution of the veridical shape indices. For the stimuli used in this experiment, $\overline{\Delta s}_{\max} = 1.57$. If the adjusted shape indices would be random, the expected averaged shape index error would be $\overline{\Delta s}_{\text{random}} = 0.69$ (dotted gray line). Note, however, that uniformly distributed adjusted shape indices do not necessarily mean that the corresponding gauge figure settings are random.

shape of the adjusted surface differs from that of the veridical surface. An error of Δs of 0.5 occurs, for example, if a cylindrical surface is misjudged to be saddle-like (and vice versa) or if a convex (or concave) surface is misjudged to be cylindrical. While Δs is below 0.5 for 91.5 % of the surface locations in the opaque base condition, this is only true for 78.4 % and 67.7 % in the hollow and massive transparent base conditions (see Figure 4.23 left). Accordingly, the average Δs is higher for massive transparent objects ($\overline{\Delta s}_{\text{Trns:Mass:Full}} = 0.439$) than for hollow transparent ($\overline{\Delta s}_{\text{Trns:Holl:Full}} = 0.319$) or opaque objects ($\overline{\Delta s}_{\text{Trns:Mass:Full}} = 0.180$, Figure 4.23 right). This pattern of result is similar to that found for the normal error Δn. In contrast to the depth error Δd, the shape index error Δs is only slightly reduced if the surface reconstruction is based on the mean adjusted normals $\bar{\hat{n}}_i$ instead of the individually adjusted normals $\hat{n}_{ik}$.

Analogous to the shape index error, we defined the curvedness error as the absolute difference between the adjusted and the veridical curvedness ($\Delta c_i = |\hat{c}_i - c_i|$, where i denotes a specific vertex of the reconstructed surface). The results show that Δc is very similar in all massive and hollow transparent stimulus conditions (see Figure 4.24a). Differences mainly occur in the opaque stimulus conditions. While Δc is about 30 % smaller in the opaque base condition than in the transparent conditions, the relative strength of the influence in

the different opaque cue conditions is similar to that indicated by the normal error Δn (cf. Figure 4.15b). Like Δs, Δc is only slightly reduced if the adjusted surfaces are reconstructed from the mean adjusted normals.

In order to detect any systematic under- or overestimation of the curvedness, we further analyzed the curvedness bias ($\mathrm{B}c_i = \hat{c}_i - c_i$, where i denotes a specific vertex of the reconstructed surface). As Figure 4.24b shows, the mean Bc is negative for all stimulus conditions. This indicates that subjects generally underestimated the objects' curvedness. Although this underestimation occurs for both opaque and transparent objects, there are differences between the stimulus conditions. For example, the bias is more pronounced for hollow transparent than for opaque and massive transparent objects. If surfaces are reconstructed from the mean adjusted normals, Bc becomes more negative. This is, however, not surprising, because averaging the adjusted surface normals can attenuate individual surface features adjusted by different subjects which in turn decreases the overall curvedness of the adjusted surface.

4.6.4.8 Spatial error distribution

The evaluation of the gauge figure settings and the quantities derived from them, such as the shape index, has so far been carried out mainly in a summarized form, i.e. by averaging the corresponding parameters over the various objects shown in the experiment and/or their individual points of measurement and the subjects. In contrast, we will now analyze the spatial distribution of these parameters at the level of individual objects. This approach makes it possible to identify specific image areas in which the subjects make larger errors than in others. In particular, increased systematic deviations from the veridical values in some regions can provide clues as to how the subjects used certain image information for their settings.

With respect to the normal error Δn, it can be seen that its spatial distribution across the surface is relatively inhomogeneous, especially for some of the massive transparent objects. As an example, Figure 4.25a shows the distribution for the object "Obj2" (cf. Figure 4.12). While Δn tends to be smaller in the peripheral areas of this object (cf. Figure 4.18), there is a ring-shaped area in the middle of the object where Δn is considerably higher ($\Delta n_{\max} = 102.4°$). Near the center of this ring-like area the normal error is again much smaller ($\overline{\Delta n} \approx 45°$). A separate analysis for the three subjects who adjusted the gauge figures in this case reveals that only two of them show this specific distribution of Δn (see Figure 4.25b, subjects "AEMA" and "DUUN"). The distribution for the remaining subject is much more homogeneous and the overall error level smaller (see Figure 4.25b, subject "RARA"). Apparently, the high absolute precision variance that we found for massive transparent objects (cf. Figure 4.16b) has a systematic cause itself, namely that different subjects perceived the shape of the object differently.

To investigate how the subjects perceived the shape of object "Obj2", we compared the individual surfaces reconstructed from their settings (see Figure 4.26 left column). The results suggest that the two subjects with an inhomogeneous spatial distribution of Δn (subjects

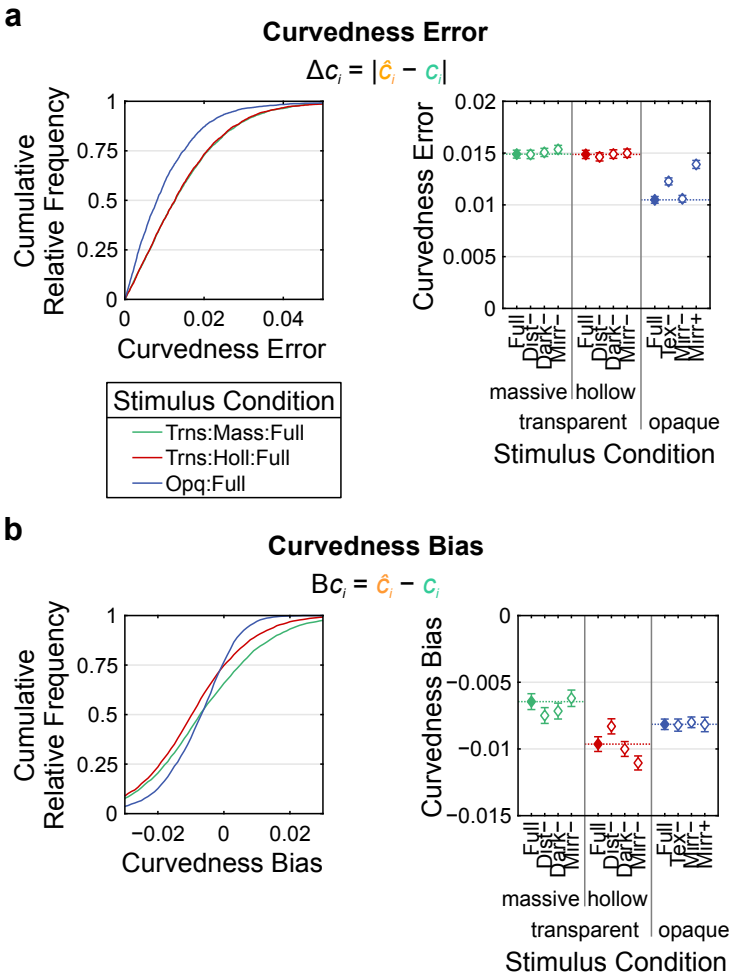

Figure 4.24: Evaluation of the curvedness error and the curvedness bias. **(a)** Analysis of the curvedness error $\Delta c_i = |\hat{c}_i - c_i|$ with $\hat{c}_i$ being the local curvedness of the adjusted, c_i being the local curvedness of the veridical surface and i a specific vertex of the reconstructed surface. The left side shows the cumulative frequency distribution of Δc for the transparent and opaque base conditions, pooled across all objects, points of measurements and subjects. The right side shows Δc (±95 %-CI) for all stimulus conditions, averaged across all objects, points of measurements and subjects. **(b)** Analysis of the curvedness bias $\mathrm{B}c_i = \hat{c}_i - c_i$. The left side shows the cumulative frequency distribution of $\mathrm{B}c_i$ for the transparent and opaque base conditions, pooled across all objects, points of measurements and subjects. The right side shows $\mathrm{B}c$ (±95 %-CI) for all stimulus conditions, averaged across all objects, points of measurements and subjects.

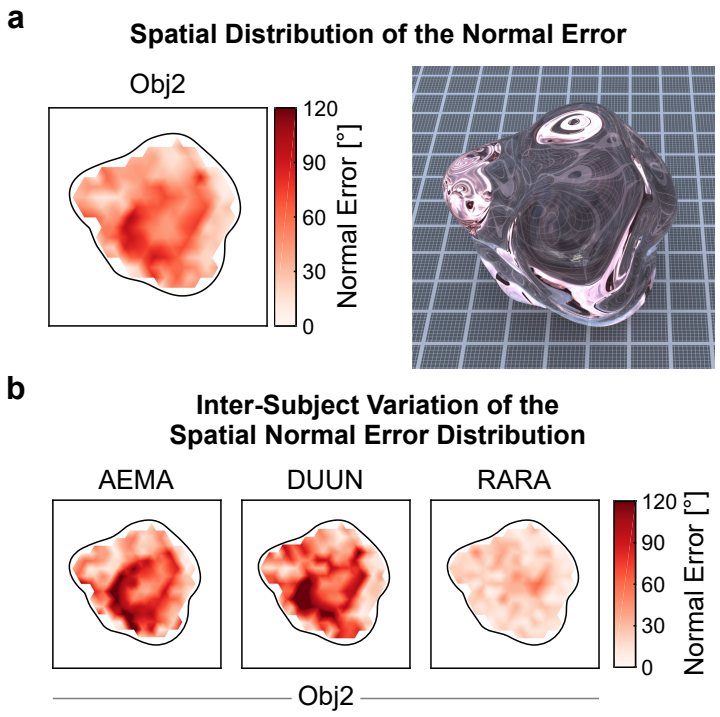

Figure 4.25: Analysis of the spatial distribution of the normal error Δn. **(a)** Spatial distribution of the normal error Δn for the object "Obj2" in the stimulus condition "Trns:Mass:Full", averaged across all subjects (left diagram). The right side shows the corresponding stimulus image (right eye only, trimmed). **(b)** Spatial distribution of the normal error Δn shown separately for the three subjects ("AEMA", "DUUN", "RARA") that were presented with the object "Obj2" in the stimulus condition "Trns:Mass:Full".

"AEMA" and "DUUN") saw a surface that has an indentation near the object's center, whereas the remaining subject with a spatially more homogeneous Δn (subject "RARA") has correctly perceived a bulge in the corresponding area. In the other cue conditions some of the subjects have also seen an indentation, except in the cue condition in which background distortions were missing ("Dist−"). Systematic misinterpretations of the local shape also occurred with other objects. This is, for example, clearly visible for the object "Obj3" (see Figure 4.26 right column). Here, two of three subjects have also erroneously judged a convex surface area in the middle of the object to be concave. In contrast to object "Obj2", this systematic misjudgment also occurred when background distortions were missing ("Dist−"), whereas it did not occur when darkening was omitted ("Dark−"). We will discuss possible explanations for these systematic misinterpretations of the surface shape in more detail in the discussion (see Section 4.6.5).

4.6.4.9 Material and massiveness ratings

Figure 4.27 shows the results of the follow-up survey. In both the opaque and the hollow transparent base condition ("Opq:Full" and "Trns:Holl:Full", respectively) the material of the example object was correctly identified by all subjects. In the massive transparent base condition ("Trns:Mass:Full") this is only true for 90.5 % of the subjects. In the transparent case, the largest number of misclassifications occurred for the massive object without mirror reflections ("Trns:Mass:Mirr−") and the hollow object without background distortions ("Trns:Holl:Dist−"). In these cases, 47.6 % and 42.9 % of the subjects incorrectly judged the sample object to be opaque. In the opaque case, only the completely reflecting object (cue condition "Mirr+") was not perceived as opaque by all subjects. 23.8 % have incorrectly classified the corresponding object as being transparent.

In the three base conditions, the massiveness was correctly identified by 47.6 % of the subjects for the massive transparent, by 76.2 % for the hollow transparent and by 95.2 % for the opaque object. In the cue conditions of the hollow transparent case, the proportion of correct massiveness estimates was either equal to or slightly lower than in the corresponding base condition. The lowest value was found in the cue condition without reflections ("Mirr−"). Only 52.4 % of the subjects correctly recognized that object as being hollow. The completely reflecting opaque example object (cue condition "Mirr+") was incorrectly identified as being hollow by 33.3 % of the subjects.

4.6.5 Discussion

In this experiment, we presented subjects with stereoscopic images of randomly shaped transparent objects, either hollow or massive, and asked them to indicate the orientation of the normal at various surface points. We varied the availability of three potential shape cues

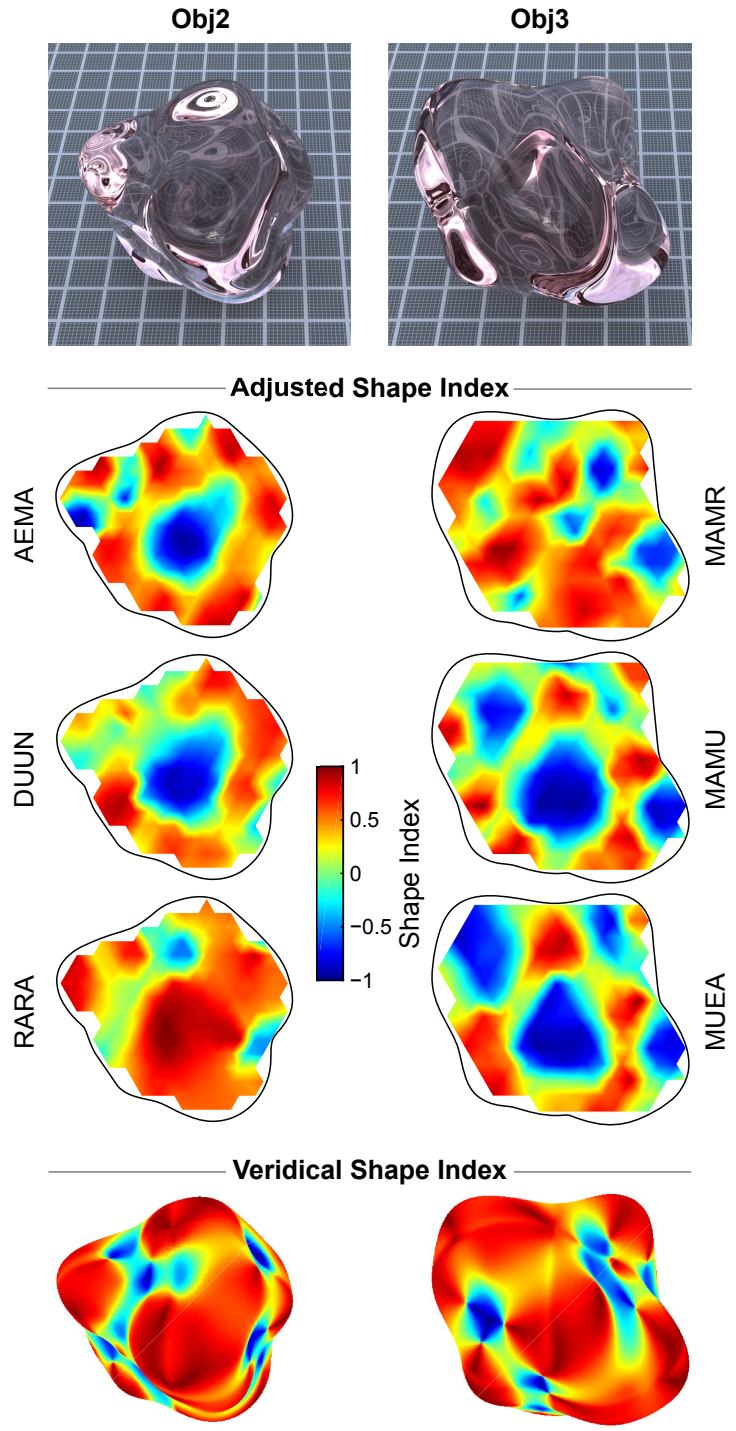

Figure 4.26: Analysis of inter-subject differences in the perceived shape of two massive transparent objects (left column object "Obj2", right column object "Obj3"). From top to bottom, this figure shows the respective stimulus images (right eye only, trimmed), the shape indices of the surfaces that were reconstructed from the gauge figure settings of the respective subjects (subject code next to it), and the veridical shape indices of the object meshes. Both the reconstructed and the veridical surfaces are shown with the same perspective projection as the stimulus images shown at the top.

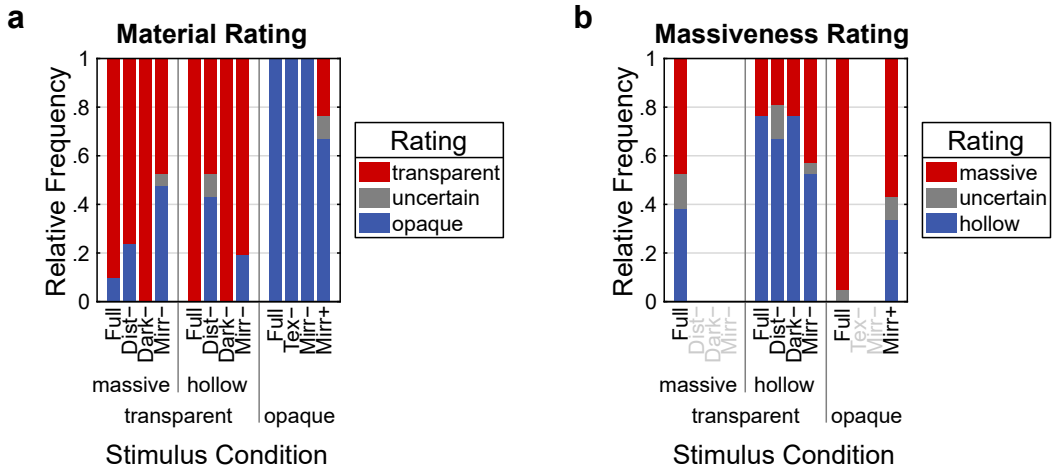

Figure 4.27: Results of the follow-up survey, in which the subjects were asked to indicate the material and the massiveness of an example object shown in different stimulus conditions. Massiveness ratings were only performed by subjects of the second group, to whom hollow objects were shown during the experiment. The ratings are based on printed copies of stimulus images. Furthermore, all ratings refer to the same example object and not to the objects actually seen by the subjects in the experiment. **(a)** Stacked bar plot showing the relative frequency of the material ratings for each stimulus condition, averaged across all subjects. **(b)** Stacked bar plot showing the relative frequency of the massiveness ratings for each stimulus condition, averaged across all subjects of the second group.

(background distortions due to refraction, darkening due to absorption, and mirror images due to specular reflections) by altering either scene and material properties or the image generation. For comparison, we also presented subjects with opaque objects and also varied the availability of corresponding shape cues.

Our computational analysis revealed that the potential shape cues in the transparent case are particularly complex. The main reason is that they are not only related to the shape of the object but also depend on properties of the whole scene. It is therefore not surprising that the shape of transparent objects was judged less accurately than that of opaque ones. On average, the errors made for transparent objects were approximately twice as large as those made for opaque objects. This suggests that the visual system processes shape-related image information differently in the transparent case than in the opaque case. We also found that hollow transparent objects were perceived considerably more accurately than massive transparent objects. This result is also not surprising, as our analysis has shown that massive objects often lead to much larger background distortions than hollow objects of identical shape. From a computational point of view, the estimation of shape therefore appears to be much more complex in such situations.

For both massive and transparent objects, shape perception was influenced by the three manipulated potential shape cues. Depending on whether the objects were solid or hollow, some of the potential shape cues had opposite effects. For massive transparent objects, both the inclusion of specular reflections and absorption enhanced shape perception, whereas for

thin-walled objects of same outer shape, absorption had almost no and specular reflections even negative effects. The effect of background distortions was very small for hollow and negative for massive transparent objects. These results indicate that in the transparent case, the particular influence of an image regularity appears to depend more on the specific situation than in the opaque case. A possible explanation of this finding is that the visual system can use the image regularities for shape perception only, if they remain in certain bounds. An interesting question for further research would be to examine more explicitly how the accuracy of shape perception depends on the specific manifestation of the different image regularities, like for example on the degree of optical distortion.

The outer contour of an object seems to play a similar role for shape perception in both the opaque and the transparent case. The closer to the rim of an object, the more similar the error levels of both material classes are. This is not surprising because the outer contour of an object and thus the shape information provided by the corresponding edge in the image does not depend on the material. With increasing distance from the rim, the contribution of other shape cues appears to become more dominant. At least in the transparent case, however, these cues seem to be less reliable than the contour.

Although the specific role that an image regularity plays for perceiving the shape of transparent objects appears to depend on a complex interplay of properties of the object itself and its surround, our results provide a first insight into how different image regularities are processed by the visual system. However, the interpretation of the results is by no means trivial. This applies not only to those cases in which a certain image regularity had a positive effect on shape perception, but also to those cases in which no effect was visible or shape perception was negatively influenced. In the following, we will briefly discuss these different patterns of results and the interpretations they are compatible with. In particular, there is evidence that some image regularities were used in a way that is not suitable for estimating the shape of transparent objects.

As a first example, we consider cases where cues had *no effect* on shape perception. This was the case with mirror images caused by opaque objects and absorption-induced darkening caused by hollow transparent objects. This appears to indicate that these two cues were not used for shape perception. However, it is also possible that these image regularities actually served as a shape cue but that their influence was not noticeable because they provide shape information that is consistent with the information already provided by other shape cues.

If the presence of an image regularity has a *positive effect* on shape perception, the interpretation may appear less equivocal. However, even in this case there are several ways to interpret the results. This can be illustrated by considering the darkening caused by absorption that had a positive influence on the perception of massive transparent objects. One interpretation is that the visual system actually used the thickness information provided by the darkening to estimate the shape. An alternative explanation is that the darkening has only an indirect

positive effect by enhancing the influence of another shape cue. For example, the chromaticity and brightness changes caused by absorption might have reduced the visibility of higher-order reflections and increased the visibility of the particularly informative first-order reflections.

The presence of background distortions with massive transparent objects and mirror images with hollow transparent objects had a clearly *negative effect* on shape perception. Again, there are various possible interpretations for these observations. First, the two image regularities might not serve as shape cues themselves, but might only have an indirect (negative) influence on other shape cues. For example, the strong background distortions of massive transparent objects could have made it difficult to detect the mirror images, which themselves had a clearly positive influence on shape perception. This explanation is not implausible because background distortions and mirror images are usually reflected in the image in a similar way. Both lead to spatially varying, direction-dependent magnifications and compressions of the background or the mirrored environment. If two such regularities are superimposed in the image, using one of them as a shape cue (in this case the mirror images) could be impeded.

Conversely, a negative influence of an image regularity on shape perception does not necessarily mean that it is not used by the visual system as a shape cue. The visual system might refer to the regularity, but not in a way that is appropriate to estimate the shape correctly. One reason for this could be that the image regularity exists in a form that the visual system cannot process in a meaningful way. As outlined above, the visual system might, for example, be able to use optical distortions of individual refractive surfaces as a shape cue, but not the much more complex distortions of objects with multiple refractions. If the same mechanisms are nevertheless also used in more complex situations, this could lead to errors and negative effects on shape perception. Apparently, such an erroneous use of image information is preceded by what can be regarded as a misinterpretation of the information available in the image. In the current example, for instance, distortions caused by multiple refractions might have been misinterpreted as distortions of a single refractive surface. Our results provide some evidence that such confusions do not have to be limited to the same material category. In our local analyses of the adjusted surface shapes, we described several cases in which systematic, spatially limited deviations from the veridical shape occurred. In three of four cue conditions of the objects “Obj2” and “Obj3”, two out of three subjects have erroneously perceived a convex bulge in the middle of the objects as a concave indentation (cf. Figure 4.26). In both cases, the deviating settings are compatible with the interpretation that the subjects have at least partially interpreted the available image information as if it had been caused by opaque objects.

As a first example we consider the object “Obj2”. Our results suggest that the local shape errors can be attributed to the influence of background distortions, because such errors did not occur in situations without distortions (i.e. cue condition “Dist−”). A possible explanation

for the systematic shape errors observed in this case is that the background distortions of the transparent object have been at least partially misinterpreted as the distorted texture of an opaque object. This would not be implausible because optical distortions of the background and shape-induced variations of the texture of opaque objects are hard to discern in the image. Both lead to spatially varying, direction-dependent magnifications or compressions, in one case of the background, in the other case of the surface texture. If object "Obj2" is presented as a massive transparent object, there is a ring-shaped area around its center along which there are strong directional distortions of the background (see Figure 4.28 left). In the case of textured opaque objects, such an image regularity would essentially be compatible with two interpretations. The area enclosed by the distortions could be a concave indentation or a convex bulge (see Figure 4.28 right). From this perspective, at least two of the three subjects misinterpreted the background distortions as an opaque shape cue and interpreted the distortion pattern as being caused by an indentation. This interpretation of the results is further supported by the fact that the systematic shape errors also occurred in the cue condition "Mirr−", in which almost half of the subjects erroneously said that an opaque object is shown. Since the conscious material impression indicates an opaque material, it is at least not implausible that in this and the other cue conditions the visual system erroneously uses mechanisms suitable for perceiving the shape of opaque objects. In the present case, the perception of an concave indentation could have been further supported by the darkening of the background. In the opaque case, such a darkening likely occurs in strongly concave and shaded areas of the surface. Note that in the present case, the darkening was mainly caused by the background being shadowed by the transparent object itself and not by the absorption of the object material. It is therefore also a good example of the ambiguity of the darkening information, which we mentioned earlier.

A misinterpretation of background darkening as shading of an opaque object could also be responsible for the systematic shape errors we found with object "Obj3" (see Figure 4.29). Here, two out of three subjects have in three of four cue conditions also erroneously perceived convex bulges in the middle of the object as concave indentations. In this area, the background has been darkened by absorption. The systematic shape errors did not occur if the object had no absorption that darkened the background (cue condition "Dark−").

The potential misinterpretations of image information outlined so far are not necessarily the only ones that have occurred. For both object "Obj2" and object "Obj3" it would also be possible that parts of the bright and/or strongly saturated areas with total reflections have been misinterpreted as mirror images or highlights caused by ordinary surface reflections. Some of these total reflections are located exactly at the border of the area where the systematic shape errors occurred (cf. Figure 4.28 left and Figure 4.29 left). Mirror images or highlights with a similar appearance usually occur in strongly curved, convex areas and thus indirectly support the interpretation of a concave indentation.

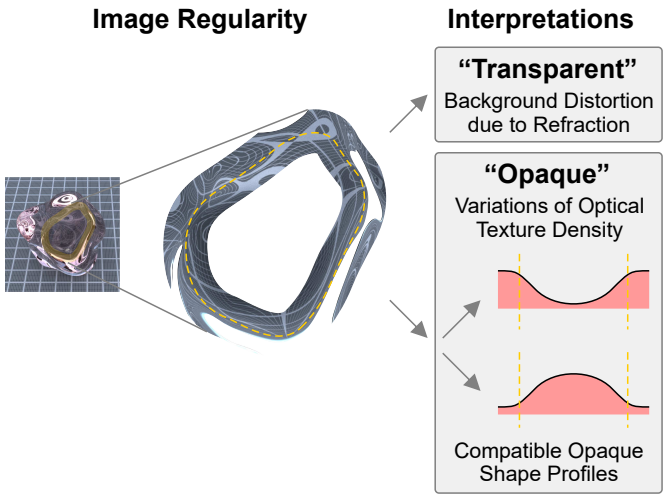

Figure 4.28: Possible misinterpretation of background distortions of a transparent object as shape-induced variations of the texture density of an opaque object. The left side shows a massive transparent object used in the experiment (“Obj2”, right eye only, trimmed). In the center of the object a ring-shaped area can be seen, along which the background is distorted (yellow-colored). Next to the stimulus, this image regularity is shown in isolation. In the case of textured opaque objects, such an image regularity would be expected at image regions where the inclination of the surface to the observer (i.e. its slant) is particularly high (see e.g. Fleming, Holtmann-Rice, & Bülthoff, 2011a). Without further information, however, the local orientation of the surface (i.e. its tilt) is ambiguous. The surface along the ring-shaped distorted area could therefore be inclined either inwards or outwards. In the first case, the area enclosed by the distortions would be a concave indentation, in the second case a convex bulge.

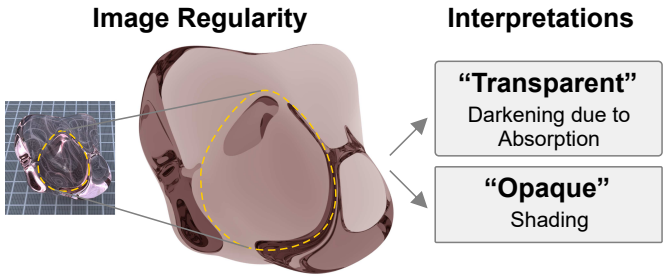

Figure 4.29: Possible misinterpretation of the darkening of a transparent object caused by absorption as shading of an opaque object. The left side shows a massive transparent object used in the experiment (“Obj3”, right eye only, trimmed). In the center of the object the background is markedly darkened due to absorption (yellow dashed area). Next to the stimulus, this image regularity is shown in isolation. If the visual system is unable to correctly identify the absorption-induced darkening as such, it might be misinterpreted as the shading of an opaque object. In the opaque case, such a darkening would be expected, for example, in strongly concave and correspondingly shaded surface areas.

Whether the convex bulge set by the remaining subjects was based on an opaque or transparent interpretation of the image information, cannot easily be determined. Based on the current results it is also difficult to judge whether these misinterpretations are merely a spatially limited phenomenon or even a general characteristic of the perception of transparent materials. To answer this question, more detailed predictions for potential opaque misinterpretations would be required. For the complex shaped objects with multiple reflections and refractions used in this experiment, such predictions can be ambiguous and can partly overlap with those for transparent interpretations.

4.7 General discussion

In this work, we dealt both theoretically and empirically with the visual perception of the shape of transparent objects. On the theoretical level, we analyzed how the shape of transparent objects is reflected in properties of the (retinal) image. In particular, we considered several image regularities associated with shape that are specific for transparent objects. These include optical distortions of the background caused by light refraction, changes in chromaticity and brightness caused by absorption, and distorted mirror images of the environment caused by specular reflections at each surface that separates spatial regions with different refraction indices. Our computational analyses showed that the relationship between these image regularities and shape are often substantially more complex than in the opaque case. Furthermore the analyses showed that the common problem that cues interact with each other and cannot be considered in isolation occurs in a particularly pronounced form with transparent objects.

The substantial differences of shape-related information that is available in the images of transparent and opaque objects strongly suggest that shape perception works differently for objects from these two material classes. This raised the question how well the visual system can recognize the shape of transparent objects at all and how well this is possible in comparison to opaque objects. A further question was, whether the potential shape cues that were identified in our theoretical analysis actually play a role in shape perception. To investigate these questions, we conducted an experiment in which we used a gauge figure task to measure the accuracy and precision of shape perception depending on the availability of potential shape cues. Our results show that the subjects' settings in the transparent case were both less accurate and less precise than in the opaque case. Furthermore, the influence of individual image regularities in the transparent case was sometimes opposite, depending on whether they originated from massive or hollow objects.

These observed differences between solid and hollow objects are a consequence of the fact that for transparent objects not only the material and the visible part of the outer shape are crucial for shape perception, but also the surfaces that are not directly visible and the

properties of the interior. The interior could in principle be of almost arbitrary complexity. Thus, the massive and hollow objects used in the experiment must be considered as just two exemplars from the set of all possible objects that, given a specific outer shape, can be defined by varying the wall thickness and the number of enclosed surfaces, and *not as* representatives of two disjunct object classes.

The present results also indicate that the influence that image regularities have on shape perception can vary greatly within the set of possible objects. Our finding that some transparent objects were misinterpreted as opaque objects suggests that the differences related to the object type are partly due to the fact that in some cases strategies that are only suited for shape perception in the opaque case were applied to transparent objects. It seems that a general model of shape perception in the transparent case must also include a model about how the object type influences the use of specific image regularities. A preliminary step for solving this general problem would be to identify for each object type the ranges of relevant object and scene parameters in which the regularities can be used at all. A pragmatic approach for determining such parameter ranges would be to determine empirically for each situation whether and in which parameter ranges shape perception works with acceptable accuracy. Based on such data, more abstract principles could be identified that determine these parameter ranges. On the basis of our theoretical analysis of background distortions, it seems plausible to assume that the absence of intersections of light rays along their path is such an abstract principle.

Although our results indicate an influence of some image regularities on the perception of shape, it is difficult to tell in which way this happens. This applies in particular to the question, whether the respective image regularities were actually used as a shape cue. Especially with massive transparent objects, our results provide some evidence that the image regularities caused by refraction and absorption were partially misinterpreted as opaque shape cues. In addition, some image regularities might have had only a moderating effect on other shape cues. Therefore, a primary goal of subsequent studies should be to obtain further information about how certain image regularities influence the shape perception of transparent objects. For this purpose, cases in which the perceived shape deviates markedly from the veridical shape are particularly diagnostic. Different types of deviations promise specific insights into the mechanisms underlying shape perception, but also require different methodological approaches.

One type of deviations results from unsystematic errors. This means that the shape parameters set by the subjects scatter randomly around the veridical shape parameters. In the cases we investigated, the size of such deviations varied considerably depending on the image regularities that were present. Although different degrees of unsystematic errors are partly compatible with several interpretations, a deeper analysis of the pattern of such errors can nevertheless help to obtain further information about the respective role of the image

regularities under scrutiny. If an image regularity is actually used as a shape cue, then it is to be expected that the size of the unsystematic error depends on the amount of usable information the cue provides. Variations in the size of random errors can then be indicative about the degree of uncertainty in estimating the shape based on this cue. Accordingly, intentional changes in the amount of information should lead to corresponding changes in the error level. This could be achieved by varying the background texture, the illumination or the absorption properties of the transparent material. For example, the contrast of the background texture could be altered or more localized light sources could be used instead of a complex ambient illumination. If the systematic errors are reduced in situations that presumably have a higher information density, this would provide further evidence that the corresponding image regularity is actually used as a shape cue. In addition, such manipulations are also a proper means of identifying the parameter ranges within which the individual image regularities can be used as a shape cue by the visual system.

Not only unsystematic errors are diagnostic for the mechanisms underlying shape perception, but also systematic errors. Examples are the systematic misinterpretation of the local surface shape of two massive transparent objects that we found in the present investigation. A possible interpretation of these deviations is that the visual system did not use the image regularities properly, but that image regularities caused by refraction and absorption were misinterpreted as originating from opaque objects and were accordingly used in a way that would only be appropriate to infer the shape of opaque objects. In future work, the validity of this explanation should be examined in more detail. A first interesting question is, how regular these misinterpretations are and whether the perceived local shapes can actually be divided into only two disjunct categories. Using a larger number of subjects might help to obtain more reliable information about the frequency of those misinterpretations. In addition, an attempt should be made to model the misinterpretations more precisely. This could be based on a more detailed analysis of the information and regularities available in the image. On this basis, more specific hypotheses could be generated as to which aspects of the image could be responsible for the misinterpretations. Ideally, it should be possible to make precise predictions as to which shape an image regularity would be compatible with if it were incorrectly interpreted as an opaque shape cue. These hypotheses could then be verified by deliberately manipulating these regularities in the image. In the cases observed in the present investigation, for example, the plausibility of an opaque interpretation of the background distortions could be increased by omitting all other transparency-specific image regularities. The proportion of subjects with systematic misinterpretations of the shape should then increase. An opposite strategy would be to remove image aspects that supposedly trigger the misinterpretation. The systematic misinterpretations should then no longer occur. Alternatively, the stimuli could be modified in order to contain additional cues to the material. Such cues can arise, for example, from movements of the object, parts of the scene, or the observer. The dynamics and their regu-

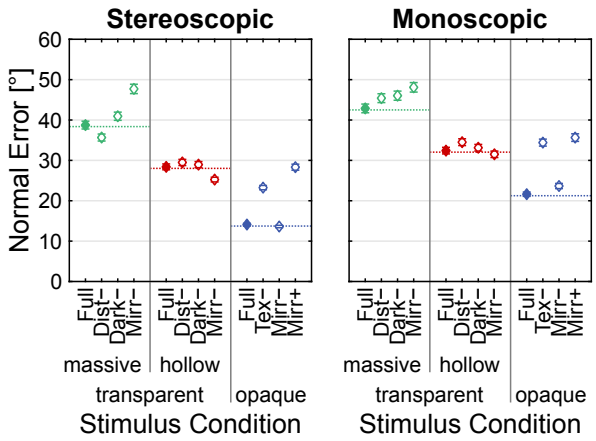

Figure 4.30: Comparison of the results obtained in the current experiment with stereoscopic stimuli (left) and preliminary results obtained for a replication of the experiment with monoscopic stimuli (right). Each diagram shows the angular normal error (±95 %-CI) for each stimulus condition, averaged across all objects, points of measurement and subjects.

larities in the image not only are known to influence the activation of transparency-specific mechanisms (Kawabe & Kogovšek, 2017; Kawabe et al., 2015; Tamura, Higashi, & Nakauchi, 2018), but could also provide cues to the shape itself (e.g. Ben-Ezra & Nayar, 2003). From a theoretical point of view, it therefore appears promising to also consider moving stimuli and to analyze their shape-related image regularities in more detail.

In the present experiment, we used stereoscopic stimuli to simulate natural viewing conditions. It is an interesting question, whether and how binocular disparity that is available in this case influences the processing of individual image regularities and shape perception in the transparent case. A natural assumption seems to be that it contributes positively to shape perception as in the case of opaque objects (Doorschot, Kappers, & Koenderink, 2001). However, it is also possible that the multiple disparity patterns related to different image regularities within mirror images or background distortions are too complex to be useful for the visual system. It is even possible that these complex disparity patterns have a negative effect on shape perception by hampering the evaluation of other shape cues. To investigate the actual influence of binocular disparity in the transparent case, it would be interesting to compare the results obtained with stereoscopic and monoscopic stimuli. We currently replicate the present experiment with such monoscopic stimuli, i.e. we present the same images to both eyes. Preliminary results of this study suggest that the overall effect of this change of viewing mode on shape perception is not very large, in the sense that the general pattern of results is rather similar (see Figure 4.30). However, there are also noticeable changes in the direction and strength of some effects. For instance, including the distortion cue in the case of massive transparent objects decreases the normal error slightly in the monoscopic case, whereas it leads to an increase of the error in the stereoscopic case.

A major challenge in investigating the perception of the shape of transparent objects is that there are numerous interactions between different image regularities associated with shape. Due to these interactions, it was difficult in our experiment to decide whether an image regularity had an effect because it was actually used as a shape cue or merely because it had an indirect positive or negative influence on other image regularities. Although such interactions complicate conclusions about individual image regularities, they also open up further methodological options for investigating the role that these image regularities play. The basic idea is to manipulate different aspects of certain image regularities separately from each other. This includes, on the one hand, the aspect associated with the shape (shape-related part) and, on the other hand, the aspect indirectly influencing other shape cues (moderating part). A major advantage of this approach is that shape-related and moderating parts can largely be described and analyzed at the level of image generation. The hypothesis that a certain image regularity has only an indirect effect, without serving as a shape cue itself, could then be tested empirically by removing its shape-related part, but preserving its moderating part. Using the example of absorption-induced darkening, this would mean that the higher order reflections are darkened in the image, but with a constant degree of darkness across the stimulus, so that it does not provide systematic information about material thickness.

Conclusions Our empirical results have shown that subjects can at least approximately recognize the shape of transparent objects. We have analyzed several image regularities related to transparent objects with respect to their suitability as shape cues. All of these potential cues had, at least in some cases, a noticeable effect on the accuracy and precision of shape perception. However, the shape perception of transparent objects was substantially worse than that of opaque objects. In addition, the performance depended on the actual object type and was considerably better for thin-walled hollow objects than for massive ones. We discussed several strategies how the analysis of local and global errors may be used to derive more specific hypotheses about the role played by individual image regularities in shape perception.

Appendix

4.A Example scene parameters

The example scenes and objects were created with the 3D modeling software Blender (Blender Foundation, 2015) and the actual images were rendered with the physically based Cycles renderer (Blender Foundation, 2015). Both modeling and rendering was performed in RGB color space.

The mesh of the example object was based on an icosahedron that was subdivided six times. The resulting icosphere consisted of 81920 triangular faces and was adjusted to a diameter of 100 mm. The icosphere was deformed by translating its vertices along their normal direction (modifier "Displace" with parameters "Direction" = "Normal", "Midlevel" = 0.5, and "Strength" = 1). The amount of displacement was determined by the intensity of three-dimensional Perlin noise (texture "Cloud" with parameters "Noise Basis" = "Original Perlin", "Size" = 1, and "Depth" = 0 and options "Grayscale" and "Soft" selected)

The example object was either opaque or transparent. In the transparent case, the object was either massive or hollow. In the hollow case (see Figure 4.9b right, Figure 4.11c bottom), it had a wall thickness of 1 mm and was created by eroding its interior without changing its outer shape (modifier "Solidify" with parameters "Thickness" = 2 and "Offset" = −1 and options "Even Thickness" and "High Quality Normals" selected).

In Figure 4.1a left, the object is made of a plastic-like material with a smooth textured surface. To implement this material, two surface shaders were mixed (first surface shader "Diffuse BSDF" with parameters "Color" given by a texture and "Roughness" = 0; second surface shader "Glossy BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(1, 1, 1), and "Roughness" = 0). The two shaders were mixed based on the Fresnel equations ("Mix Shader" with parameter "Fac" given by a "Fresnel" node with parameter "IOR" = 1.49). The granite-like texture was based on three-dimensional Perlin noise (texture "Cloud" with parameters "Noise Basis" = "Original Perlin", "Size" = 1.5, and "Depth" = 2 and options "Grayscale" and "Soft" selected). The color of the texture ranged from RGB(.144, .096, .096) to RGB(.336, .224, .224).

In Figure 4.9a right, Figure 4.9b right, and Figure 4.11c, the object was made of acrylic glass that does not absorb light (surface shader "Glass BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(1, 1, 1), "Roughness" = 0, and "IOR" = 1.49; no volume shader). In Figure 4.1b

left, the glass was tinted red (surface shader “Glass BSDF” with parameters “BSDF” = “Sharp”, “Color” = RGB(1, 1, 1), “Roughness” = 0, and “IOR” = 1.49; volume shader “Volume Absorption” with parameters “Color” = RGB(.84, .76, .76) and “Density” = 0.03).

The object was surrounded by (and, if hollow, filled with) non-absorptive air (refractive index $R \approx 1$).

The floor below the object was either textured with a gray graph paper, a gravel pattern or consisted of a uniform gray. The gray graph paper (surface shader “Diffuse BSDF” with parameter “Color” given by a procedural texture and “Roughness” = 0) was used in Figure 4.1a left, Figure 4.1b left, Figure 4.9a right, and Figure 4.9b right. The node setup of the procedural texture consisted of two superimposed grid textures of different size (node “Brick Texture”, grid color = RGB(.12, .12, .12), grid width 1 mm and 10 mm, background color = RGB(.28, .28, .28)).

The uniform gray background color (surface shader “Diffuse BSDF” with parameters “Color” = RGB(.2, .2, .2) and “Roughness” = 0) was used in Figure 4.11c.

Image based lighting with an infinitely distant spherical emitter was used. The illumination texture was a high dynamic range image (color depth 16 bit/channel) of a natural daylight outdoor scene with a partly cloudy sky (Yimm & Bell, 2008).

A perspective camera was placed at a distance of 400 mm to the center of the object (vertical field of view 44.10°).

The scenes were rendered as high dynamic range images with the Cycles renderer (image size 1040 × 1040 px, 2048 samples/px, color depth 16 bit/channel) and subsequently tonemapped within Blender to low dynamic range images (color depth 8 bit/channel) as well as gamma corrected ($\gamma = 1.2$).

4.B Shape index and curvedness

The shape index and the curvedness describe the type of local surface shape and the strength of local curvature, respectively (Koenderink & van Doorn, 1992).

The shape index s is defined as

$$s = \frac{2}{\pi} \arctan\left(\frac{K_{\min} + K_{\max}}{K_{\min} - K_{\max}}\right), \tag{4.1}$$

where $K_{\min}$ and $K_{\max}$ denote the principal curvatures of the surface. These principal curvatures describe the maximum and minimum curvature at a given surface location. The directions of $K_{\min}$ and $K_{\max}$ are orthogonal to each other. The shape index s describes the type of local surface shape on the continuum from spherical concave ($s = -1$, $K_{\min} = K_{\max} < 0$, “Cup”) to saddle-like ($s = 0$, $K_{\min} = -K_{\max}$, “Saddle”) to spherical convex shapes ($s = +1$, $K_{\min} = K_{\max} > 0$, “Cap”).

The curvedness c is defined as

$$c = \sqrt{\frac{K_{\text{min}}^2 + K_{\text{max}}^2}{2}} \tag{4.2}$$

and describes the strength of local curvature independently from the type of surface shape. The higher the curvedness c, the more accentuated the local surface shape is. In contrast to the shape index s, the value of the curvedness c depends on the actual size of an object.

4.C Analysis of absorption-induced darkening

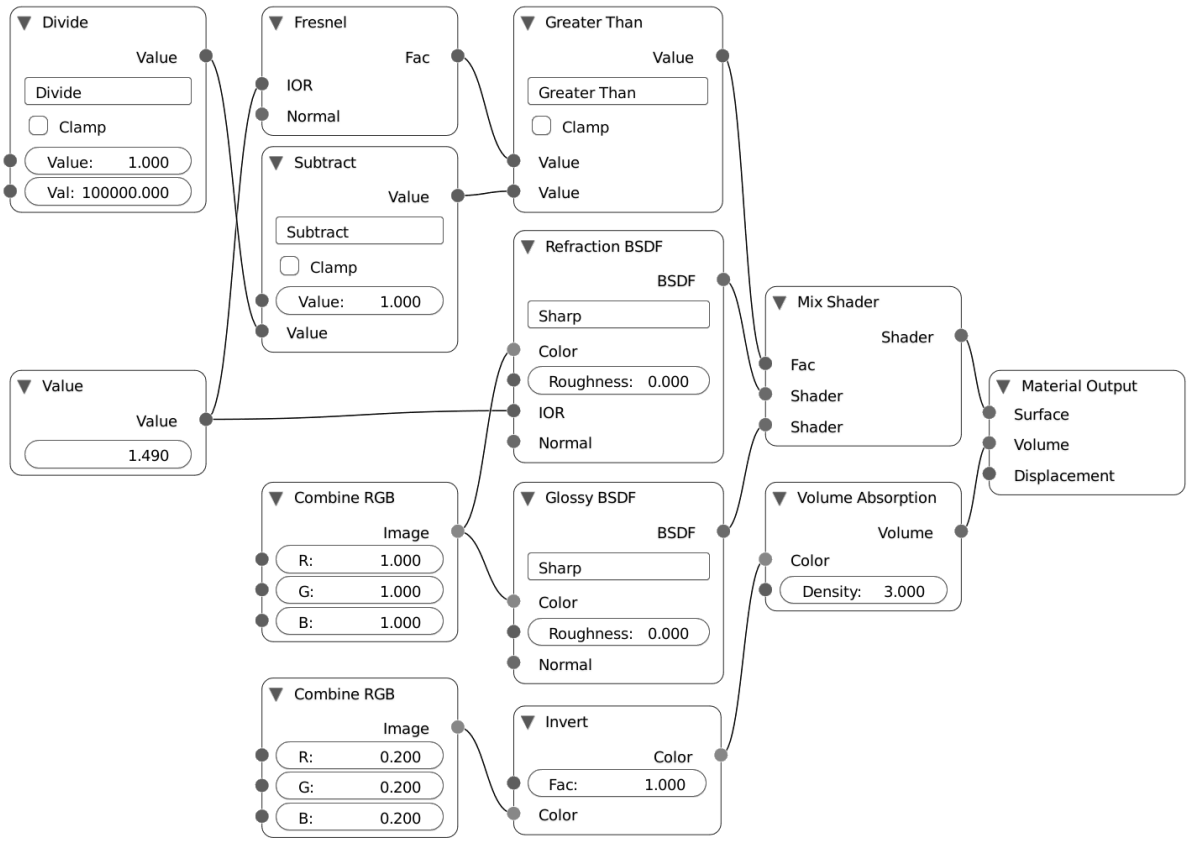

Figure 4.C.1: Blender node setup defining the Cycles object material used to obtain the absorption-induced darkening in the image. See Section 4.3.1 for details. The illustration is based on a screenshot of the graphical user interface of Blender (Blender Foundation, 2015).

4.D Mirror images of different order

To isolate mirror images of different order, we changed the material properties of the example object (cf. Appendix 4.A) so that different parts of its surface had different transmission and reflection properties. For this purpose, the object mesh was split into two parts. The cutting plane was orthogonal to the observation direction and was located in the center of gravity

of the icosphere from which the object mesh was derived. After the split, the solid object consisted of two meshes (front and back surface), the hollow object of four meshes (outer front, inner front, inner back and outer back surface).

To isolate the first-order mirror images, the front surface was changed to a black plastic-like material. This was achieved by using two surface shaders (first surface shader "Diffuse BSDF" with parameters "Color" = RGB(0, 0, 0) "Roughness" = 0); second surface shader "Glossy BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(1, 1, 1) and "Roughness" = 0) that were mixed based on the Fresnel equations ("Mix Shader" with parameter "Fac" given by a "Fresnel" node with parameter "IOR" = 1.49). Since the light of first-order reflections interacts with the front surface of an object only, the mirror image is the same for the massive and the hollow object.

To isolate the second-order mirror images, the outer front surface has been set so that is refracts light but does not reflect it. This was achieved by using two surface shaders (first surface shader "Refraction BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(1, 1, 1) "Roughness" = 0) and "IOR" = 1.49; second surface shader "Glossy BSDF" with parameters "BSDF" = "Sharp", "Color" = RGB(0, 0, 0) and "Roughness" = 0) that were mixed based on the Fresnel equations ("Mix Shader" with parameter "Fac" given by a "Fresnel" node with parameter "IOR" = 1.49). Note that, despite the absence of reflections, the transmitted light was influenced by Fresnel effects. For the massive object, the back surface of the object was changed to the black plastic-like material (see shader description above). For the hollow object, the inner front surface was changed to the black plastic-like material.

To isolate the third-order mirror image of the hollow object, both its outer and inner front surface were changed to refract light but not to reflect it (see shader description above). The inner back surface of the object was changed to consist of the black plastic-like material (see shader description above).

To isolate the fourth-order mirror image of the hollow object, only its outer back surface was adjusted to be made of the black plastic-like material, while all other surfaces were changed to refract light but not to reflect it.

The scenes were rendered as high dynamic range images with the Cycles renderer (Blender Foundation, 2015; image size 1040×1040 px, 512 samples/px, color depth 16 bit/channel) and subsequently tonemapped within Blender to low dynamic range images (color depth 8 bit/channel) as well as gamma corrected ($\gamma = 1.2$).

This procedure to isolate mirror images of different order is only a rough approximation. It is based on the simplifying assumption that mirror images of different order are separated from each other by the middle plane of the object and that mirror images occurring on the rear half of the object are not influenced by reflections at the front half of the object.

4.E Variance decomposition of the normals

Variance decomposition was performed as described by Mardia and Jupp (2000, pp.163). The total spherical variance $V_{\mathrm{T}i}$ describes the total variation of the adjusted normals $\hat{n}_{ik}$ about the corresponding veridical normal n_i and is defined as

$$V_{\mathrm{T}i} = V_{\mathrm{P}i} + V_{\mathrm{A}i}, \tag{4.3}$$

where $V_{\mathrm{T}i} \in [0, 1]$ with $V_{\mathrm{T}i} = 0$ if the adjusted normals all point in the veridical direction and $V_{\mathrm{T}i} = 1$ if the adjusted normals were equally distributed in all directions.

The precision variance $V_{\mathrm{P}i}$ describes the variation of the adjusted normals $\hat{n}_{ik}$ about their mean $\overline{\hat{n}}_i$ and is defined as

$$V_{\mathrm{P}i} = 1 - \overline{R}_i, \tag{4.4}$$

where

$$\overline{R}_i = ||\overline{\hat{n}}_i|| \tag{4.5}$$

and

$$\overline{\hat{n}}_i = \frac{1}{m} \sum_{k=1}^{m} \hat{n}_{ik}. \tag{4.6}$$

Since for each object and stimulus condition, the gauge figure at a specific point of measurement was adjusted by three different subjects, m equals 3.

The accuracy variance $V_{\mathrm{A}i}$ describes the variation of the mean setting $\overline{\hat{n}}_i$ about the veridical normal n_i and is defined as

$$V_{\mathrm{A}i} = \overline{R}_i - \overline{C}_i \tag{4.7}$$

with

$$\overline{C}_i = \frac{1}{m} \sum_{k=1}^{m} n_i^{\mathrm{T}} \hat{n}_{ik}, \tag{4.8}$$

where n_i^{T} denotes the transpose of n_i.

4.F Transformation of reconstructed surface from image space to wold space

To convert the reconstructed surfaces from the image space into the world space, both their vertices and their normals were transformed. During the transformation, the normals were not considered as location-independent directions but as vectors originating from a specific vertex. To this end, we treated the tips of the normal vectors as regular points and transformed them in the same way as the vertices.

In the first step, we translated the vertices v and the normal tips t so that the zero point of the coordinate system corresponds to the location of the virtual camera. In addition, we reversed the z axis of the coordinate system:

$$v' = \begin{pmatrix} v_x - (w/2) \\ v_y - (h/2) \\ (-v_z) + (d/r) \end{pmatrix}, \tag{4.9}$$

$$t' = \begin{pmatrix} t_x - (w/2) \\ t_y - (h/2) \\ (-t_z) + (d/r) \end{pmatrix}. \tag{4.10}$$

Here, $w = 839\,\text{px}$ and $h = 1200\,\text{px}$ are the width and the height of the stimulus image, respectively, $d = 400\,\text{mm}$ corresponds to the viewing distance of the camera and $r = 0.27\,\text{mm/px}$ corresponds to the resolution of the LCD screen used in the experiment.

In the next step, we defined camera ray vectors c_v and c_t that point to the locations where the vertices and the normal tips are projected on the image plane, respectively:

$$c_v = \begin{pmatrix} v'_x \\ v'_y \\ d/r \end{pmatrix}, \tag{4.11}$$

$$c_t = \begin{pmatrix} t'_x \\ t'_y \\ d/r \end{pmatrix}. \tag{4.12}$$

The world space coordinates of the vertices and the normal tips were then calculated by adjusting the length of the camera ray vectors to the z component of the transformed vertex and normal tip vectors:

$$v'' = (c_v/||c_v||) \times v'_z, \tag{4.13}$$

$$t'' = (c_t/||c_t||) \times t'_z. \tag{4.14}$$

After this, the z axis was reversed and the zero point of the coordinate system was changed to correspond to the center of the object. In addition, the unit was changed to millimeters:

$$v''' = \begin{pmatrix} v''_x \times r \\ v''_y \times r \\ (-(v''_z \times r)) + d \end{pmatrix}, \tag{4.15}$$

$$t''' = \begin{pmatrix} t''_x \times r \\ t''_y \times r \\ (-(t''_z \times r)) + d \end{pmatrix}. \tag{4.16}$$

In the last step, the transformed vertex and normal tips vectors were used to calculate the surface normals n in the world space:

$$n = (t''' - v''')/||(t''' - v''')||. \tag{4.17}$$

Literatur

Adato, Y., Vasilyev, Y., Ben-Shahar, O. & Zickler, T. (2007). Toward a theory of shape from specular flow. In *IEEE 11th International Conference on Computer Vision (ICCV)* (S. 1–8). Rio de Janeiro, Brasilien: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICCV.2007.4408883

Adato, Y., Vasilyev, Y., Zickler, T. & Ben-Shahar, O. (2010). Shape from specular flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(11), 2054–2070. https://doi.org/10.1109/TPAMI.2010.126

Anderson, B. L. (2015). The perceptual representation of transparency, lightness, and gloss. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization*. Oxford, England: Oxford University Press.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–193. https://doi.org/10.1037/h0054663

Barrow, H. G. & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, *17*, 75–116. https://doi.org/10.1016/0004-3702(81)90021-7

Beck, J. (1978). Additive and subtractive color mixture in color transparency. *Perception & Psychophysics*, *23*(3), 265–267. https://doi.org/10.3758/BF03204137

Beck, J., Prazdny, K. & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics*, *35*(5), 407–422. https://doi.org/10.3758/BF03203917

Becket, W. & Badler, N. I. (1990). Imperfection for realistic image synthesis. *The Journal of Visualization and Computer Animation*, *1*(1), 26–32. https://doi.org/10.1002/vis.4340010108

Ben-Ezra, M. & Nayar, S. K. (2003). What does motion reveal about transparency? In *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV)* (Bd. 2, S. 1025–1032). Nizza, Frankreich: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICCV.2003.1238462

Bergmann, L., Schaefer, C. & Bergmann, L. (1993). *Optik*. (H. Niedrig, Hrsg.) (9. Aufl). Berlin, Deutschland: de Gruyter.

Bergmann Tiest, W. M. & Kappers, A. M. L. (2007). Haptic and visual perception of roughness. *Acta Psychologica*, *124*(2), 177–189. https://doi.org/10.1016/j.actpsy.2006.03.002

Bernhard, M., Waldner, M., Plank, P., Solteszova, V. & Viola, I. (2016). The accuracy of gauge-figure tasks in monoscopic and stereo displays. *IEEE Computer Graphics and Applications*, *36*(4), 56–66. https://doi.org/10.1109/MCG.2016.45

Blake, A. & Bülthoff, H. (1990). Does the brain know the physics of specular reflection? *Nature*, *343*(6254), 165–168. https://doi.org/10.1038/343165a0

Blake, A. & Bülthoff, H. (1991). Shape from specularities: Computation and psychophysics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *331*(1260), 237–252. https://doi.org/10.1098/rstb.1991.0012

Blake, A., Bülthoff, H. & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research*, *33*(12), 1723–1737. https://doi.org/10.1016/0042-6989(93)90037-W

Blender Foundation. (2013). Blender - A 3D modelling and rendering package (Version 2.66a) [Computerprogramm]. Abgerufen von https://www.blender.org

Blender Foundation. (2015). Blender - A 3D modelling and rendering package (Version 2.76) [Computerprogramm]. Abgerufen von https://www.blender.org

Chen, J. & Allison, R. S. (2013). Shape perception of thin transparent objects with stereoscopic viewing. *ACM Transactions on Applied Perception*, *10*(3), 1–15. https://doi.org/10.1145/2506206.2506208

Chevreul, M. E. (1855). *The principles of harmony and contrast of colors and their applications to the arts* (2. Auflage). London, England: Longman, Brown, Green, and Longmans.

Chowdhury, N. S., Marlow, P. J. & Kim, J. (2017). Translucency and the perception of shape. *Journal of Vision*, *17*(3), 17. https://doi.org/10.1167/17.3.17

Christou, C. G. & Koenderink, J. J. (1997). Light source dependence in shape from shading. *Vision Research*, *37*(11), 1441–1449. https://doi.org/10.1016/S0042-6989(96)00282-9

Cole, F., Sanik, K., DeCarlo, D., Finkelstein, A., Funkhouser, T., Rusinkiewicz, S. & Singh, M. (2009). How well do line drawings depict shape? *ACM Transactions on Graphics*, *28*(3), 1–9. https://doi.org/10.1145/1531326.1531334

Cutting, J. E. & Millard, R. T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General*, *113*(2), 198–216. https://doi.org/10.1037/0096-3445.113.2.198

De Haan, E., Erens, R. G. F. & Noest, A. J. (1995). Shape from shaded random surfaces. *Vision Research*, *35*(21), 2985–3001. https://doi.org/10.1016/0042-6989(95)00050-A

Doorschot, P. C. A., Kappers, A. M. L. & Koenderink, J. J. (2001). The combined influence of binocular disparity and shading on pictorial shape. *Perception & Psychophysics*, *63*(6), 1038–1047. https://doi.org/10.3758/BF03194522

Dövencioglu, D. N., Ben-Shahar, O., Barla, P. & Doerschner, K. (2017). Specular motion and 3D shape estimation. *Journal of Vision*, *17*(6), 3. https://doi.org/10.1167/17.6.3

Dubuisson, M.-P. & Jain, A. K. (1994). A modified Hausdorff distance for object matching. In *Proceedings of the 12th International Conference on Pattern Recognition* (Bd. 1, S. 566–568). Jerusalem, Israel: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICPR.1994.576361

Ernst, M. O. & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. https://doi.org/10.1016/j.tics.2004.02.002

Faul, F. (2017). Toward a perceptually uniform parameter space for filter transparency. *ACM Transactions on Applied Perception*, *14*(2), 1–21. https://doi.org/10.1145/3022732

Faul, F. & Ekroll, V. (2002). Psychophysical model of chromatic perceptual transparency based on substractive color mixture. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *19*(6), 1084–1095. https://doi.org/10.1364/JOSAA.19.001084

Faul, F. & Ekroll, V. (2011). On the filter approach to perceptual transparency. *Journal of Vision*, *11*(7), 1–33. https://doi.org/10.1167/11.7.7

Faul, F. & Ekroll, V. (2012). Transparent layer constancy. *Journal of Vision*, *12*(12), 1–26. https://doi.org/10.1167/12.12.7

Faul, F. & Falkenberg, C. (2015). Transparent layer constancy under changes in illumination color: Does task matter? *Vision Research*, *116*, 53–67. https://doi.org/10.1016/j.visres.2015.09.003

Feldman, J. & Singh, M. (2005). Information along contours and object boundaries. *Psychological Review*, *112*(1), 243–252. https://doi.org/10.1037/0033-295X.112.1.243

Fleming, R. W. & Bülthoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception*, *2*(3), 346–382. https://doi.org/10.1145/1077399.1077409

Fleming, R. W., Holtmann-Rice, D. & Bülthoff, H. H. (2011a). Estimation of 3D shape from image orientations. *Proceedings of the National Academy of Sciences, USA*, *108*(51), 20438–20443. https://doi.org/10.1073/pnas.1114619109

Fleming, R. W., Jäkel, F. & Maloney, L. T. (2011b). Visual perception of thick transparent materials. *Psychological Science*, *22*(6), 812–820. https://doi.org/10.1177/0956797611408734

Fleming, R. W., Torralba, A. & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, *4*(9), 798–820. https://doi.org/10.1167/4.9.10

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA, USA: MIT Press.

Gaissert, N. & Wallraven, C. (2012). Categorizing natural objects: A comparison of the visual and the haptic modalities. *Experimental Brain Research*, *216*(1), 123–134. https://doi.org/10.1007/s00221-011-2916-4

Gårding, J. (1990). Shape from texture and contour by weak isotropy. In *Proceedings of the 10th International Conference on Pattern Recognition* (Bd. 1, S. 324–330). Atlantic City, NJ, USA: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICPR.1990.118124

Gårding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision*, *2*(4), 327–350. https://doi.org/10.1007/BF00121877

Geier, J. & Hudák, M. (2011). Changing the Chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold - A psychophysical refutation. *PLoS ONE*, *6*, e26062. https://doi.org/10.1371/journal.pone.0026062

Gerardin, P., de Montalembert, M. & Mamassian, P. (2007). Shape from shading: New perspectives from the Polo Mint stimulus. *Journal of Vision*, *7*(11), 1–11. https://doi.org/10.1167/7.11.13

Gibson, J. J. (1950). *The perception of the visual world*. Oxford, England: Houghton Mifflin.

Grynberg, A. & Ward, G. (1990). *LBL Building 90 Third Floor Conference Room* [Digitales Modell].

Handbook of Kodak photographic filters. (1990). Rochester, NY, USA: Professional Photography Division, Eastman Kodak Co.

Hata, S., Saitoh, Y., Kumamura, S. & Kaida, K. (1996). Shape extraction of transparent object using genetic algorithm. In *Proceedings of the 13th International Conference on Pattern Recognition* (S. 684–688). Wien, Österreich: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICPR.1996.547652

Horn, B. K. P. (1970). *Shape from shading: A method for obtaining the shape of a smooth opaque object from one view*. Cambridge, MA, USA: Massachusetts Institute of Technology.

Horn, B. K. P. (1975). Obtaining shape from shading information. In P. Winston (Hrsg.), *The Psychology of Computer Vision* (S. 115–155). New York, NY, USA: McGraw-Hill.

Horn, B. K. P. (1990). Height and gradient from shading. *International Journal of Computer Vision*, *5*(1), 37–75. https://doi.org/10.1007/BF00056771

Ikeuchi, K. & Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, *17*(1), 141–184. https://doi.org/10.1016/0004-3702(81)90023-0

Interrante, V., Fuchs, H. & Pizer, S. (1995). Enhancing transparent skin surfaces with ridge and valley lines. In *Proceedings of the 1995 Conference on Visualization* (S. 52-59). Atlanta, GA, USA: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/VISUAL.1995.480795

Interrante, V., Fuchs, H. & Pizer, S. M. (1997). Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, *3*(2), 98–117. https://doi.org/10.1109/2945.597794

Ishihara, S. (1969). *Tests for color blindness.* Tokyo, Japan: Kanehara Shuppan Co. Ltd.

Jakob, W. (2013). Mitsuba renderer (Version 0.4.4) [Computerprogramm]. Abgerufen von https://www.mitsuba-renderer.org

Kasrai, R. & Kingdom, F. A. (2001). Precision, accuracy, and range of perceived achromatic transparency. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *18*(1), 1–11.

Kawabe, T. & Kogovšek, R. (2017). Image deformation as a cue to material category judgment. *Scientific Reports*, *7*, 44274. https://doi.org/10.1038/srep44274

Kawabe, T., Maruya, K. & Nishida, S. (2015). Perceptual transparency from image deformation. *Proceedings of the National Academy of Sciences, USA*, *112*(33), 4620–4627. https://doi.org/10.1073/pnas.1500913112

Kennedy, J. M. & Domander, R. (1985). Shape and contour: The points of maximum change are least useful for recognition. *Perception*, *14*(3), 367–370. https://doi.org/10.1068/p140367

Kersten, M. A., Stewart, A. J., Troje, N. & Ellis, R. (2006). Enhancing depth perception in translucent volumes. *IEEE Transactions on Visualization and Computer Graphics*, *12*(5), 1117–1123. https://doi.org/10.1109/TVCG.2006.139

Khang, B. G. & Zaidi, Q. (2002a). Accuracy of color scission for spectral transparencies. *Journal of Vision*, *2*(6), 451–466. https://doi.org/10.1167/2.6.3

Khang, B. G. & Zaidi, Q. (2002b). Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching. *Vision Research*, *42*(2), 211–226. https://doi.org/10.1016/S0042-6989(01)00252-8

Kim, J. & Marlow, P. J. (2016). Turning the world upside down to understand perceived transparency. *i-Perception*, *7*(5), 1–5. https://doi.org/10.1177/2041669516671566

Kleffner, D. A. & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*(1), 18–36. https://doi.org/10.3758/BF03206757

Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape? *Perception*, *13*(3), 321–330. https://doi.org/10.1068/p130321

Koenderink, J. J. (1998). Pictorial relief. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *356*(1740), 1071–1086. https://doi.org/10.1098/rsta.1998.0211

Koenderink, J. J., Kappers, A. M., Todd, J. T., Norman, J. F. & Phillips, F. (1996). Surface range and attitude probing in stereoscopically presented dynamic scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 869–878. https://doi.org/10.1037/0096-1523.22.4.869

Koenderink, J. J. & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta: International Journal of Optics*, *27*(7), 981–996. https://doi.org/10.1080/713820338

Koenderink, J. J. & van Doorn, A. J. (1992). Surface shape and curvature scales. *Image and Vision Computing*, *10*(8), 557–564. https://doi.org/10.1016/0262-8856(92)90076-F

Koenderink, J. J., van Doorn, A. J., Christou, C. & Lappin, J. S. (1996). Shape constancy in pictorial relief. *Perception*, *25*(2), 155–164. https://doi.org/10.1068/p250155

Koenderink, J. J., van Doorn, A. J. & Kappers, A. M. L. (1996). Pictorial surface attitude and local depth comparisons. *Perception & Psychophysics*, *58*(2), 163–173. https://doi.org/10.3758/BF03211873

Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*(3), 389–412. https://doi.org/10.1016/0042-6989(94)00176-M

Larson, G. W. & Shakespeare, R. (2003). *Rendering with Radiance: The art and science of lighting visualization*. Davis, CA, USA: Space & Light.

Lee, C.-H. & Rosenfeld, A. (1985). Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, *26*(2), 125–143. https://doi.org/10.1016/0004-3702(85)90026-8

Li, Y., Pizlo, Z. & Steinman, R. M. (2009). A computational model that recovers the 3D shape of an object from a single 2D retinal representation. *Vision Research*, *49*(9), 979–991. https://doi.org/10.1016/j.visres.2008.05.013

Lobay, A. & Forsyth, D. A. (2006). Shape from texture without boundaries. *International Journal of Computer Vision*, *67*(1), 71–91. https://doi.org/10.1007/s11263-006-4068-8

Malik, J. (1987). Interpreting line drawings of curved objects. *International Journal of Computer Vision*, *1*(1), 73–103. https://doi.org/10.1007/BF00128527

Malik, J. & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision*, *23*(2), 149–168. https://doi.org/10.1023/A:1007958829620

Mamassian, P. & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research*, *36*(15), 2351–2367. https://doi.org/10.1016/0042-6989(95)00286-3

Mamassian, P. & Landy, M. S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research*, *38*(18), 2817–2832. https://doi.org/10.1016/S0042-6989(97)00438-0

Mardia, K. V. & Jupp, P. E. (2000). *Directional statistics*. Chichester, NY, USA: J. Wiley.

Marr, D. (1977). Analysis of occluding contour. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *197*(1129), 441–475. https://doi.org/10.1098/rspb.1977.0080

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA, USA: W. H. Freeman.

Mausfeld, R. (2003). „Colour“ as part of the format of different perceptual primitives: The dual coding of colour. In R. Mausfeld & D. Heyer (Hrsg.), *Colour perception - Mind and the physical world* (S. 381–436). Oxford, England: Oxford University Press.

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics*, *13*(1), 59–66. https://doi.org/10.1080/00140137008931118

Mingolla, E. & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics*, *53*(3), 137–151. https://doi.org/10.1007/BF00342882

Morgenstern, Y., Murray, R. F. & Harris, L. R. (2011). The human visual system's assumption that light comes from above is weak. *Proceedings of the National Academy of Sciences, USA*, *108*(30), 12551–12553. https://doi.org/10.1073/pnas.1100794108

Morris, N. J. W. & Kutulakos, K. N. (2007). Reconstructing the surface of inhomogeneous transparent scenes by scatter-trace photography. In *IEEE 11th International Conference on Computer Vision* (S. 1–8). Rio de Janeiro, Brasilien: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICCV.2007.4408882

Morris, N. J. W. & Kutulakos, K. N. (2011). Dynamic refraction stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(8), 1518–1531. https://doi.org/10.1109/TPAMI.2011.24

Motoyoshi, I. (2010). Highlight-shading relationship as a cue for the perception of translucent and transparent materials. *Journal of Vision*, *10*(9), 1–11. https://doi.org/10.1167/10.9.6

Murase, H. (1990). Surface shape reconstruction of an undulating transparent object. In *Proceedings of the 3rd International Conference on Computer Vision* (S. 313–317). Osaka, Japan: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/ICCV.1990.139539

Murase, H. (1992). Surface shape reconstruction of a nonrigid transparent object using refraction and motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*(10), 1045–1052. https://doi.org/10.1109/34.159906

Muryy, A. A., Welchman, A. E., Blake, A. & Fleming, R. W. (2013). Specular reflections and the estimation of shape from binocular disparity. *Proceedings of the National Academy of Sciences, USA*, *110*(6), 2413–2418. https://doi.org/10.1073/pnas.1212417110

Nefs, H. T. (2008). Three-dimensional object shape from shading and contour disparities. *Journal of Vision*, *8*(11), 1–16. https://doi.org/10.1167/8.11.11

Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J. & Zielke, G. (2004a). The visual and haptic perception of natural object shape. *Perception & Psychophysics*, *66*(2), 342–351. https://doi.org/10.3758/BF03194883

Norman, J. F., Phillips, F. & Ross, H. E. (2001). Information concentration along the boundary contours of naturally shaped solid objects. *Perception*, *30*(11), 1285–1294. https://doi.org/10.1068/p3272

Norman, J. F. & Raines, S. R. (2002). The perception and discrimination of local 3-D surface structure from deforming and disparate boundary contours. *Perception & Psychophysics*, *64*(7), 1145–1159. https://doi.org/10.3758/BF03194763

Norman, J. F., Todd, J. T. & Orban, G. A. (2004b). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science*, *15*(8), 565–570. https://doi.org/10.1111/j.0956-7976.2004.00720.x

Norman, J. F., Todd, J. T. & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, *57*(5), 629–636. https://doi.org/10.3758/BF03213268

Oren, M. & Nayar, S. K. (1994). Generalization of Lambert's reflectance model. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (S. 239–246). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/192161.192213

Oren, M. & Nayar, S. K. (1997). A theory of specular surface geometry. *International Journal of Computer Vision*, *24*(2), 105–124. https://doi.org/10.1023/A:1007954719939

O'Shea, J. P., Banks, M. S. & Agrawala, M. (2008). The assumed light direction for perceiving shape from shading. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (S. 135–142). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1394281.1394306

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA, USA: MIT Press.

Panis, S., De Winter, J., Vandekerckhove, J. & Wagemans, J. (2008). Identification of everyday objects on the basis of fragmented outline versions. *Perception*, *37*(2), 271–289. https://doi.org/10.1068/p5516

Pentland, A. (1989). Shape information from shading: A theory about human perception. *Spatial Vision*, *4*(2), 165–182. https://doi.org/10.1163/156856889X00103

Phillips, F., Egan, E. J. L. & Perry, B. N. (2009). Perceptual equivalence between vision and touch is complexity dependent. *Acta Psychologica*, *132*(3), 259–266. https://doi.org/10.1016/j.actpsy.2009.07.010

Pizlo, Z. (2008). *3D shape: Its unique place in visual perception*. Cambridge, MA, USA: MIT Press.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166. https://doi.org/10.1038/331163a0

Reinhard, E. & Devlin, K. (2005). Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, *11*, 13–24. https://doi.org/10.1109/TVCG.2005.9

Richards, W. A., Koenderink, J. J. & Hoffman, D. D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America*, *4*(7), 1168–1175. https://doi.org/10.1364/JOSAA.4.001168

Ripamonti, C., Westland, S. & Da Pos, O. (2004). Conditions for perceptual transparency. *Journal of Electronic Imaging*, *13*(1), 29–35. https://doi.org/10.1117/1.1636764

Robilotto, R., Khang, B. G. & Zaidi, Q. (2002). Sensory and physical determinants of perceived achromatic transparency. *Journal of Vision*, *2*(5), 388–403. https://doi.org/10.1167/2.5.3

Rosenholtz, R. & Malik, J. (1997). Surface orientation from texture: Isotropy or homogeneity (or both)? *Vision Research*, *37*(16), 2283–2293. https://doi.org/10.1016/S0042-6989(96)00121-6

Rusinkiewicz, S. (2004). Estimating curvatures and their derivatives on triangle meshes. In *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT)* (S. 486–493). Thessaloniki, Griechenland: Institute of Electrical and Electronics Engnineers. 10.1109/TDPVT.2004.1335277

Savarese, S., Chen, M. & Perona, P. (2004a). Recovering local shape of a mirror surface from reflection of a regular grid. In T. Pajdla & J. Matas (Hrsg.), *Computer Vision - ECCV 2004* (S. 468–481). Berlin, Deutschland: Springer. https://doi.org/10.1007/978-3-540-24672-5_37

Savarese, S., Chen, M. & Perona, P. (2005). Local shape from mirror reflections. *International Journal of Computer Vision*, *64*(1), 31–67. https://doi.org/10.1007/s11263-005-1086-x

Savarese, S., Fei-Fei, L. & Perona, P. (2004b). What do reflections tell us about the shape of a mirror? In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization* (S. 115–118). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1012551.1012571

Savarese, S. & Perona, P. (2001). Local analysis for 3D reconstruction of specular surfaces. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (Bd. 2, S. 738–745). Kauai, HA, USA: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/CVPR.2001.991038

Savarese, S. & Perona, P. (2002). Local analysis for 3D reconstruction of specular surfaces - Part II. In A. Heyden, G. Sparr, M. Nielsen & P. Johansen (Hrsg.), *Computer Vision - ECCV 2002* (Bd. 2351, S. 759–774). Berlin, Deutschland: Springer. https://doi.org/10.1007/3-540-47967-8_51

Schlüter, N. & Faul, F. (2014). Are optical distortions used as a cue for material properties of thick transparent objects? *Journal of Vision*, *14*(14), 1–14. https://doi.org/10.1167/14.14.2

Schlüter, N. & Faul, F. (2016). Matching the material of transparent objects: The role of background distortions. *i-Perception*, *7*(5), 1–24. https://doi.org/10.1177/2041669516669616

Schlüter, N. & Golz, J. (2015). An illumination representation approach to the Chevreul effect. *Perception*, *44*(6), 662–678. https://doi.org/10.1177/0301006615594269

Schwärzler, M. & Wimmer, M. (2007). *Rendering imperfections: Dust, scratches, aging,...* (No. TR-186-2-07-09). Wien, Österreich: Institute of Computer Graphics and Algorithms, Vienna University of Technology. Abgerufen von https://www.cg.tuwien.ac.at/research/publications/2007/TR-186-2-07-09/

Sharan, L., Rosenholtz, R. & Adelson, E. H. (2009). Material perception: What can you see in a brief glance? *Journal of Vision*, *9*(8), 784–784. https://doi.org/10.1167/9.8.784

Sharan, L., Rosenholtz, R. & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision*, *14*(9), 12–12. https://doi.org/10.1167/14.9.12

Shih, Y., Krishnan, D., Durand, F. & Freeman, W. T. (2015). Reflection removal using ghosting cues. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (S. 3193–3201). Boston, MA, USA: Institute of Electrical and Electronics Engnineers. https://doi.org/10.1109/CVPR.2015.7298939

Singh, M. & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review*, *109*(3), 492–519. https://doi.org/10.1037/0033-295X.109.3.492

Stevens, K. A. (1980). *Surface perception from local analysis of texture and contour* (AITR-512). Cambridge, MA, USA: Massachusetts Institute of Technology. Abgerufen von https://dspace.mit.edu/handle/1721.1/6869

Stevens, K. A. (1981). The information content of texture gradients. *Biological Cybernetics*, *42*(2), 95–105. https://doi.org/10.1007/BF00336727

Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, *46*(3), 183–195. https://doi.org/10.1007/BF00336800

Stockman, A. & Sharpe, L. T. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, *40*(13), 1711–1737. https://doi.org/10.1016/S0042-6989(00)00021-3

Stockman, A., Sharpe, L. T. & Fach, C. (1999). The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches. *Vision Research*, *39*(17), 2901–2927. https://doi.org/10.1016/S0042-6989(98)00225-9

Sun, J. & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, *1*(3), 183–184. https://doi.org/10.1038/630

Tamura, H., Higashi, H. & Nakauchi, S. (2018). Dynamic visual cues for differentiating mirror and glass. *Scientific Reports*, *8*(1), 8403. https://doi.org/10.1038/s41598-018-26720-x

Taylor, A. H. & Kerr, G. P. (1941). The distribution of energy in the visible spectrum of daylight. *Journal of the Optical Society of America*, *31*(1), 3. https://doi.org/10.1364/JOSA.31.000003

Tittle, J. S., Norman, J. F., Perotti, V. J. & Phillips, F. (1998). The perception of scale-dependent and scale-independent surface structure from binocular disparity, texture, and shading. *Perception*, *27*(2), 147–166. https://doi.org/10.1068/p270147

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, *8*(3), 115–121. https://doi.org/10.1016/j.tics.2004.01.006

Todd, J. T. & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 242–255. https://doi.org/10.1037/0096-1523.13.2.242

Todd, J. T., Norman, J. F., Koenderink, J. J. & Kappers, A. M. (1997). Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception*, *26*(7), 807–822. https://doi.org/10.1068/p260807

Todd, J. T., Oomes, A. H. J., Koenderink, J. J. & Kappers, A. M. L. (2004). The perception of doubly curved surfaces from anisotropic textures. *Psychological Science*, *15*(1), 40–46. https://doi.org/10.1111/j.0963-7214.2004.01501007.x

Todd, J. T. & Thaler, L. (2010). The perception of 3D shape from texture based on directional width gradients. *Journal of Vision*, *10*(5), 1–13. https://doi.org/10.1167/10.5.17

Tse, P. U. (2002). A contour propagation approach to surface filling-in and volume formation. *Psychological Review*, *109*(1), 91–115. https://doi.org/10.1037/0033-295X.109.1.91

van Doorn, A. J., Koenderink, J. J. & Wagemans, J. (2011). Light fields and shape from shading. *Journal of Vision*, *11*(3), 1–21. https://doi.org/10.1167/11.3.21

Vrhel, M. J., Gershon, R. & Iwan, L. S. (1994). Measurement and analysis of object reflectance spectra. *Color Research & Application*, *19*(1), 4–9. https://doi.org/10.1111/j.1520-6378.1994.tb00053.x

Walter, J. (2014). *04_conference_room* [Digitales Modell]. Abgerufen von https://github.com/wahn/export_multi/tree/master/04_conference_room

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In P. Winston (Hrsg.), *The Psychology of Computer Vision* (S. 19–91). McGraw-Hill.

Wijntjes, M. (2012). Probing pictorial relief: From experimental design to surface reconstruction. *Behavior Research Methods*, *44*(1), 135–143. https://doi.org/10.3758/s13428-011-0127-3

Wijntjes, M., Vota, R. M. & Pont, S. (2015). *The perceptual integrability of 3D shape.* Posterpräsentation auf der 38. European Conference on Visual Perception (ECVP), Liverpool, England.

Witkin, A. P. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence*, *17*, 17–45. https://doi.org/10.1016/0004-3702(81)90019-9

Yimm, L. & Bell, D. (2008). *Cloudy Pier* [Digitales Bild]. Abgerufen von http://www.hdrlabs.com/sibl/archive/

Lebenslauf

Name:	Nick Schlüter
seit 2011	Wissenschaftlicher Mitarbeiter, Institut für Psychologie, Christian-Albrechts-Universität zu Kiel
2013 bis 2019	Promotion im Fach Psychologie, Institut für Psychologie, Christian-Albrechts-Universität zu Kiel
2006 bis 2011	Studium der Psychologie (Diplom), Institut für Psychologie, Christian-Albrechts-Universität zu Kiel