Seeing through transparent layers

Dicle N. Dövencioğlu

Department of Psychology,
Justus-Liebig-University Giessen, Giessen, Germany
National Magnetic Resonance Research Center,
Bilkent University, Ankara, Turkey

Andrea van Doorn

KU Leuven, Leuven, Belgium
Utrecht University, Utrecht, The Netherlands

Jan Koenderink

KU Leuven, Leuven, Belgium
Utrecht University, Utrecht, The Netherlands

Katja Doerschner

Department of Psychology, Bilkent University,
Ankara, Turkey
National Magnetic Resonance Research Center,
Bilkent University, Ankara, Turkey
Department of Psychology,
Justus-Liebig-University Giessen, Giessen, Germany

The human visual system is remarkably good at decomposing local and global deformations in the flow of visual information into different perceptual layers, a critical ability for daily tasks such as driving through rain or fog or catching that evasive trout. In these scenarios, changes in the visual information might be due to a deforming object or deformations due to a transparent medium, such as structured glass or water, or a combination of these. How does the visual system use image deformations to make sense of layering due to transparent materials? We used eidolons to investigate equivalence classes for perceptually similar transparent layers. We created a stimulus space for perceptual equivalents of a fiducial scene by systematically varying the local disarray parameters *reach* and *grain*. This disarray in eidolon space leads to distinct impressions of transparency, specifically, high reach and grain values vividly resemble water whereas smaller grain values appear diffuse like structured glass. We asked observers to adjust image deformations so that the objects in the scene looked like they were seen (a) under water, (b) behind haze, or (c) behind structured glass. Observers adjusted image deformation parameters by moving the mouse horizontally (grain) and vertically (reach). For two conditions, water and glass, we observed high intraobserver consistency: responses were not random. Responses yielded a concentrated equivalence class for water and structured glass.

Introduction

Humans can expertly distinguish multiple layers in the same retinal location. This is a crucial ability when interacting with 3D objects in our environment because we are surrounded by a see-through medium, the atmosphere, which is not always uniformly clear. For example, we can effectively navigate while driving through heavy fog, hard rain, or a heat haze by separating these atmospheric effects from the optic flow caused by our own motion through the objects. This ability requires the human visual system to split visual information into different layers, such as a transparent medium and an object in the background.

Light entering a transparent medium changes its direction depending on the angle of incidence and the refraction index of the medium. This causes distortions in the image. For instance, when light waves pass through an air to water boundary, the direction of light changes depending on the refractive indices of these two mediums (Snell's law). The direction of light, hence the distortions in the image, depends on the water's surface shape, (e.g., amplitude of the waves on the water's surface; Figure 1a). Moreover, the various physical causes of light refraction introduce many types of layering in visual information. Atmospheric refraction includes phenomena such as rain, fog, or haze. For
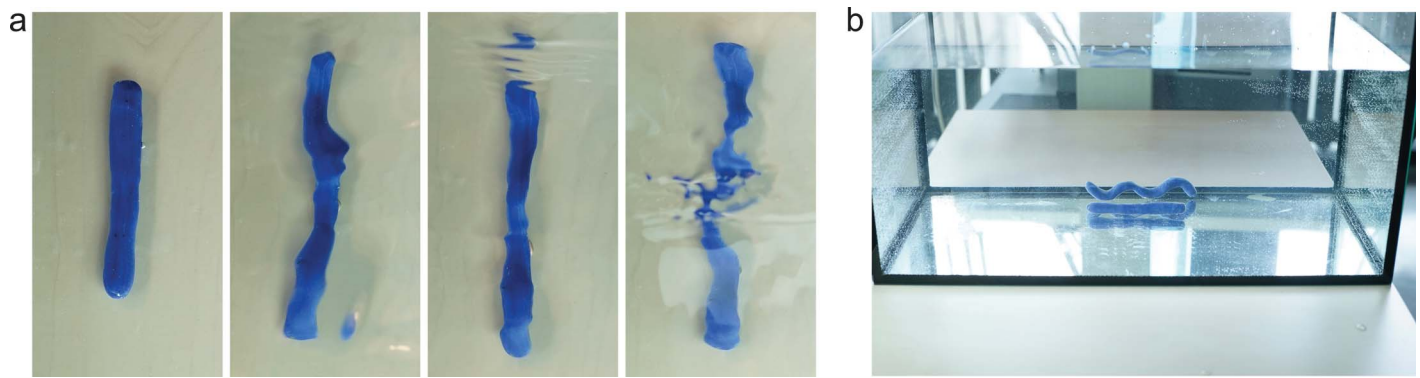
Figure 1. Refraction of light. (a) A piece of blue modeling clay is photographed under water, when the water is still (first image) and when it is wavy (remaining three images). Photographs show shape distortions and phase breaks due to the ripples on the water, making it impossible to recover the veridical shape of the modeling clay. (b) Photographing set-up viewed from the side reveals two different shapes of modeling clay inside a full water tank.

example, the third image in Figure 1a might as well be mistaken for a transparent steam layer. It can also be caused by single scattering of light: on a clear day, we can see objects in large distances, so that the sharpness of contours remain the same but contrast might change (known as airlight or Koschmieder's law). Or multiple light scattering can be caused by the molecules in the medium, such as a glass of diluted milk (Kubelka Munk theory). In the case of multiple scattering, the medium might become translucent by completely obscuring the object behind (Koenderink & van Doorn, 2001). These distinct impressions of transparency each cause characteristic disarray in images but only a small part of these many physical causes of layering is studied in the scope of visual perception.

From our own experience we know that complex visual scenes are perceptually split into multiple causal layers so that we perceive the shape and material of objects and surfaces, and the prevailing illumination in a scene. For instance, we recognize shadows as a separate layer and we do not trip over them while walking. Evidence for perceptual layer decomposition lies in our ability to interpret 3D shape, which requires the ability to distinguish between shading, shadows, and reflectance of a surface in static scenes (Zhou & Baker, 1996; Schofield, Rock, Sun, Jiang, & Georgeson, 2010; Dövencioğlu, Welchman, & Schofield, 2013; for a review, see Kingdom, 2011). Transparency perception is investigated in numerous studies in the scope of decomposing reflectance and illumination layers; these studies establish classical examples where a background texture seen through a uniform transparent layer has reduced contrast (Metelli, 1970; Anderson, 1997). In these examples, sharpness of contours remains unchanged and the layer affects only the contrast. In 3D shape-from-shading tasks with textured surfaces, the coherence of first order (local mean luminance) and second order (local luminance amplitude) local luminance values can account for the

perception of luminance and reflectance related changes (Schofield, Hesse, Rock, & Georgeson, 2006). When local luminance cues are aligned congruently, this is seen as changes in the shading gradient of a corrugated surface; whereas incongruent alignment of these cues is interpreted as the reflectance changes on a flat surface (e.g., painted stripes). The human visual system is sensitive to these cues (Schofield & Georgeson, 1999) and can perceptually learn to benefit from them in order to decompose an illumination layer from a change in surface reflectance (Dövencioğlu et al., 2013). In these elaborate findings about local changes in luminance values due to transparency, there are no comments on the potential geometric aspect of local changes in the image. How about the transparent layers, such as clear water, where the contrast remains mostly the same but the contours change? The transparent materials in daily life are mostly complex: they are thick and not uniform and the traditional photometric approaches so far do not cover the geometric distortions due to the refractive properties of transparent layers.

Transparency perception in complex scenes involving a see-through material is very different than the classically studied phenomenon; it requires the visual system to recognize the material, detect image distortions, and causally attribute those distortions to the object or the transparent medium. A couple of recent studies report global cues in images to recognize glass from mirror (Kim & Marlow, 2016; Tamura, Higashi, & Nakauchi, 2018). Fleming, Jäkel, and Maloney (2011) present compelling evidence that image distortions convey information about transparent material's intrinsic properties. Human observers estimate the refractive index of the material to make judgments about a transparent blob's thickness and the authors suggest that a mechanism similar to shape-from-texture might be involved in perceiving compressions and magnifications of the background texture due to a

transparent medium. Following studies also report that the human visual system is sensitive to image distortions due to a transparent material but argue that participants depend on image similarities, such as specular highlight shape, while matching refractive indices of thick transparent blobs (Schlüter & Faul, 2014, 2016). In another study, Kawabe, Maruya, and Nishida (2015) use spatial deformations coupled with dynamic deformations to inquire the strength of observers' perceived water layer ratings. The authors report a specific band (0.13–1.03 c/°) of spatial deformations that yield a percept of a transparent water layer when coupled with temporal distortions. These studies imply that the visual system uses image distortion cues to make judgments about thick transparent materials such as water and similarly clear liquids. Although these are appealing material examples, they constitute only a small portion of complex transparent layers.

We encounter much richer examples of transparent layers in daily life, and although they all convey some type of image distortions in the optic field, some (e.g., water and structured glass) differ considerably. Often in naturalistic examples, transparent layers carry both geometric deformations, such as seeing pebbles through clear water, and photometric deformations, such as reduced contrast due to absorbed light when seeing objects through haze or fog. In most cases, neither type of distortion is uniform throughout the image, in the sense that the impressions vary locally, sometimes due to the changing thickness of the transparent medium. Yet, the visual system readily perceives fine differences in transparent materials; for example, it is not even a challenge to distinguish multiple transparent materials when we see a cold glass of ice cubes and water with sweat beads on the glass' surface. What are the image distortions that the human visual system uses to identify these different types of transparencies?

Recently, Koenderink, Valsecchi, van Doorn, Wagemans, and Gegenfurtner (2017) presented an advantageous tool to apply image distortions locally. With the eidolon factory, one can deform an image while keeping the local disarray uniform throughout, plus control the disarray by using two simple parameters. This tool also enables us to introduce photometric distortions by changing the *coherence* in the image, thus edge quality and the contrast in the image disarray can be controlled independently. Within the eidolon factory one can parameterize image distortions locally in a way that matches our understanding of the receptive fields of the human vision. Eidolons allow us to break the image deformation magnitude into two parameters: with the *reach* parameter we control the intensity independently from the *grain* parameter that characterizes the locality of deformations. Moreover, the tool provides the conceptual freedom to define *eidolons*, an

equivalence class of appearance such that perceptually similar stimuli are eidolons; that is, they fall in the same class even if they are not identical. As an example for our specific purpose, looking into a 1-m deep pool and a 1.5-m deep pool produce slightly different images but they are likely to be perceived in the same way so these fall in the same equivalence class. The importance of this equivalence class concept manifests itself in the wide range of results in our brainstorming exercise in which observers list all the transparent materials they can think of.

Here, we present a new type of transparency that is created by local image disarray and allows us to generate materials that appear transparent. Since it is not possible to cover the wide diversity of transparent materials, we will focus on three exemplars for each of the matter states: glass, water, and haze. We first explore these most common transparency examples to parametrically identify their equivalence classes with 2D shapes defined by contours only. Next we use images of 3D objects with different surface colors and reflectances. We first ask, when the transparent layer's boundary is not visible and there is no texture in the background, whether it is still possible to induce the percept of an object seen through a transparent layer. Since local image deformations are most effective when edges are present, this condition mostly relies on the object boundary. Next, we occlude the object boundary as well to explore perceived transparency via image distortions on surface properties only. We demonstrate that, even in the absence of the object boundary, it is possible to generate perceived transparency in static images and we describe these parametrically. We investigate the role of surface reflectance with four different objects. Finally, we make an attempt at predicting these equivalence classes for perceived transparency in rigid and nonrigid complex scenes.

Brainstorming

Purpose

The few transparent material examples in the literature are focused around water and similar clear materials (Fleming et al., 2011; Kawabe et al., 2015), and haze (Kawabe & Kogovšek, 2017), but there are more examples for common transparent materials. This brainstorming session was performed to explore the rich lexicon for transparent layers in daily language, in anticipation that common words would serve as a first model to equivalence classes of transparency in complex scenes, which is a relatively new field of study.

Procedure

Prior to the experiments, in a 10-min brainstorming session, observers were given a pen and paper and asked to write down as many transparent layers as they can think of. Observers wrote detailed descriptive lists of transparent layers in German or Turkish, which were then translated into English. Throughout the translations, we picked the first meaning or the most common synonym among the English words that corresponded to the original German and Turkish words.

Observers

Fourteen naive observers participated in Experiment 1, sixteen different observers participated in Experiment 2, and 10 different observers completed Experiment 3. In total, all 40 observers completed a brainstorming session prior to the specific experiments. Participants were students at the Justus-Liebig-University Giessen, Germany (30 women, 10 men; mean age = 24.4 ± 2.5), with normal or corrected-to-normal vision. Observers gave written informed consent and they were paid 6 Euros per hour for their participation. Experimental procedures adhered to the principles put forth by the Declaration of Helsinki.

Results and discussion

Results of the brainstorming session are presented as a word cloud in Figure 2. The combined list of transparent materials included 668 items. We manually assigned a category; for instance, word list items such as most liquids, ocean, saliva, rain, and ice were in the category water; glass(es), contact lenses, windshield, window, glass bottle, magnifying glass, and diamond were in category glass. In total, 24.7% of the items belonged to the glass category. This was followed by plastic (18.7%), water (16%), air (7.6%), textile (3.7%), and paper (2.7%); the remaining items did not fall into any of these categories. Overall, 171 words were not categorized (26.5%; e.g., fingernails, fly wings, ghosts, magnetic fields, Harry Potter's invisibility cloak as in Rowling, 2014). Note that the word cloud in Figure 2 depicts accumulated words and not the related categories; that is, window, glass, and glasses are individually represented in the cloud but they all fall into the glass category in the above description.

The most frequent categories resulting from this session are somewhat in agreement with our exploratory experimental conditions of glass and water. The second frequent category in our findings, plastic, suggests a different type of transparent layer. Although

Figure 2. Brainstorming session before the experiments. Word cloud representing results of all observers during the brainstorming session prior to the experiments. Higher frequency words are indicated with larger fonts (e.g., glass, water, window, air, plastic). Color is used to increase readability.

plastic is physically separate from glass, optically they can be identical. Since we aim to isolate image distortion properties for transparent layers, it is more important that each layer has characteristic optical traits, rather than identical physical properties. Hence, we add the next frequent category, air, and include haze as a third experimental condition. Agreeing with the three states of the matter that exist under normal conditions, we carry on to study glass (solid), water (liquid), and haze (gas) in the following experiments. Finally, the richly detailed examples in each category confirm that conceptualizing transparent layers in terms of equivalence classes is more effective than, for example, searching for instances of water.

Experiment 1: Generative shapes

Motivation

We first wanted to see whether it is possible to induce perceived transparent layers in 2D shapes defined by luminance contrast and separated by different contour shapes (Figure 3). Participants adjusted image deformations by controlling reach and grain parameters on luminance defined eidolons of the generative shapes. Here, we report results for water and glass trials, results
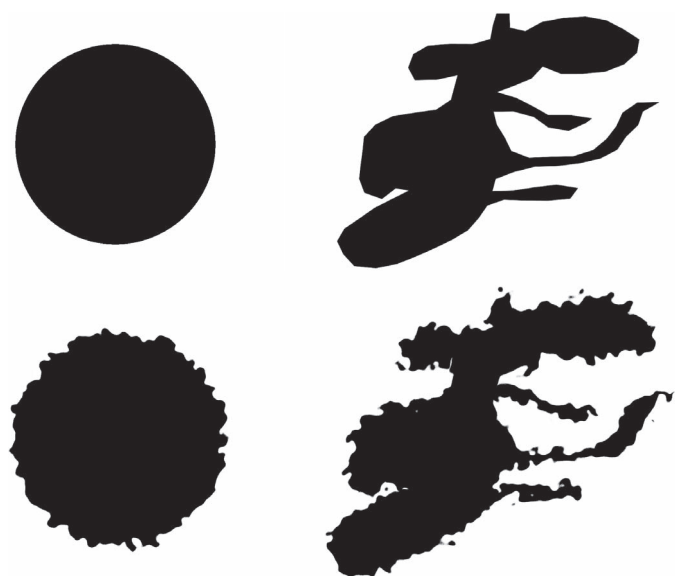
Figure 3. Motivation for the stimuli used in Experiment 1. Two fiducial images (top row) are manipulated the same way with the eidolon factory (reach = 8, grain = 8 pixels) and the resulting eidolons are shown in the bottom row.

relating to haze trials can be found in Supplementary Appendix B (Supplementary Appendix Figure B1a).

Stimuli: Fiducial images and their eidolons

We used binary luminance defined shapes as fiducial images to create stimuli for Experiment 1 (Figure 4a). In Experiment 1, fiducial images were based on a novel set of base shapes created by segmenting and combin-

ing different shapes from the MPEG-7 data set (Latecki, Lakamper, & Eckhardt, 2000; Morgenstern, Schmidt, & Fleming, 2017). Images were created as filled polygons with patch function in MATLAB, and luminance of the polygons were set to low contrast binary values (background = 0, figure = 65). Images were 600×600 pixels in Experiment 1. One may wonder, why not use a basic convex shape instead of these? The effect of distortions differ vastly depending on the shape. To illustrate these effects, we give two examples of eidolons in Figure 4. We apply the same image distortions to a simple convex shape (left column) and another shape with more boundary structure (right column). As this example depicts, the distortions are more pronounced in the second shape, and even within this shape, distortions have a more dramatic effect on the limbs compared to the convex body. In other words, the shape with limbs is maximally susceptible to distortions. Thus, we chose to include simple 2D shapes with different shapes of limbs, as can be seen in Figure 4a.

Note that the fiducial images were not shown in the experiments, we used them to create eidolons that were used as stimuli (Koenderink et al., 2017). We produced a set of controlled image deformations to the fiducial images. First, images were scale decomposed and a local spatial disarray generated by Gaussian random fields was applied to all scales; afterward, scales were combined again to construct an eidolon (i.e., a "fuzzy doppelgänger,") of the fiducial image. The local disarray is parameterized by way of two variables: The *grain* parameter is the width of the blurring kernel, which controls the rate of change in deformation as a function of local distance. The *reach* parameter denotes the intensity of the local disarray, (i.e., the amount of
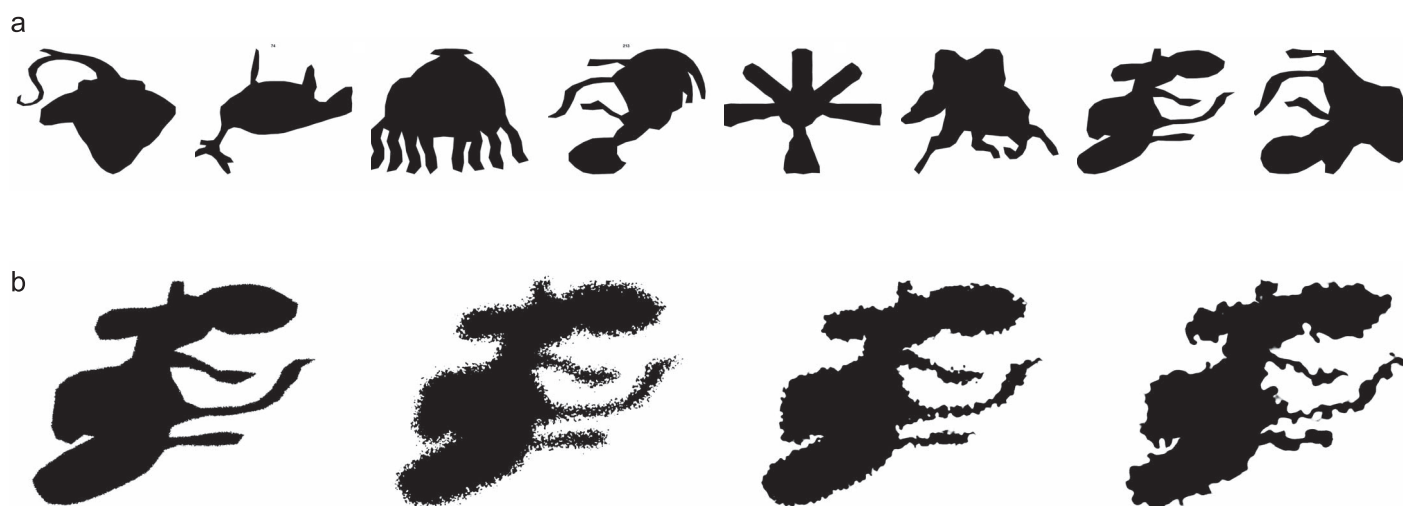
Figure 4. Examples of stimuli used in Experiment 1. (a) Eight fiducial images based on the MPEG-7 database. (b) Example eidolons of the eighth fiducial image show possible reach (R) and grain (G) adjustments with lowest parameter values used (leftmost, R = 1, G = 1), small grain and high reach values (second left, R = 8, G = 2), intermediate (third, R = 5, G = 5), and the highest parameter values used in Experiment 1 (rightmost, R = 10, G = 10).

	Experiment 1	Experiment 2		Experiment 3	
	$C = 1$	$C = 1$	$C = 0$	$C = 1$	$C = 0$
Water	812 [8,8]	1,433 [9,7]	1,377 [9,8]	1,329 [7,8]	1,290 [8,8]
Haze	**916** [5,8]	**2,449** [5,5]	**2,332** [7,7]	**1,742** [5,4]	**1,804** [5,7]
Glass	870 [2,9]	2,400 [4,8]	2,166 [4,9]	1,049 [2,9]	930 [2,10]

Table 1. L_1 distance sums for all experiments. Shown are the L_1 distance values for all trials for each layer condition (rows) in separate experiments for coherent ($C = 1$) and incoherent ($C = 0$) conditions (columns). For each condition, the smallest sum is shown next to the centroid values in brackets. The highest value in each column is bold. The centroids indicate grain and reach values, respectively, in pixel units.

displacement for a given pixel). A third parameter, *coherence*, controls whether the disarray at different scale levels is synchronized; a coherent disarray has aligned dislocations for coarse and fine structures in the image. Further details of the concept of eidolons can be found in the original paper by Koenderink et al. (2017).

For each fiducial image we created a set of 100 eidolons (1–10 pixels for reach and grain) by using the eidolon factory in MATLAB (MathWorks, Natick, MA; 2017a, https://github.com/gestaltrevision/ Eidolon). In Experiment 1, all image disarray was coherent.

Procedure

We used an adjustment method to see how observers varied image distortions for transparent layers of different nature. MATLAB Psychophysics Toolbox (Kleiner et al., 2007) was used to present stimuli on a FlexScan EV2750 monitor (EIZO, Hakusan, Ishikawa, Japan) at a resolution of $2,560 \times 1,440$ pixels. In each trial, observers saw an eidolon through an aperture (Experiments 1 and 2, hard edged, radius = 300 pixels) and above a title to indicate one of the three conditions: "Make the object look like it is seen (a) under water, (b) behind haze, (c) behind structured glass." Note that the instructions were given in German, the German word "Milchglas" was used to describe the structured glass in all experiments. Stimuli remained on screen until the observer indicated a decision by pressing the space bar. Observers were seated 50 cm from the screen, they adjusted image deformation parameters by moving the mouse horizontally (grain) and vertically (reach), which steered a white dot that was always visible on the screen.

In Experiment 1, we used eight shapes (Figure 4) for which participants adjusted image disarray for three layers (water, haze, structured glass) for two repetitions, resulting in a blocked session of 48 trials. In Experiment 1, during half of the trials, grain and reach values increased as the cursor moved rightward and upward, respectively. For the other half of trials, controls were reversed: observers had to move the cursor leftward and downward to increase grain and

reach values, respectively. This manipulation was done to see whether observers remembered the cursor's screen position instead of focusing on the image distortions, but the results were similar, so we kept the controls straightforward in Experiments 2 and 3. All trials were interleaved in one block in Experiment 1.

Analysis and data exclusion

We look at the individual mean values for reach and grain parameters separately and we report the group means in this paper. Further, we pool across participants for the reach and grain adjustments in each experiment and present these in normalized histograms. After we calculate the histogram for 10 bins (number of steps in stimulus range = 10), we divide each bin count by the maximum number of trials in a bin. In this way, the normalized histogram gives the frequency of a stimulus value instead of the raw trial count. We use a nonparametric Wilcoxon signed rank test to establish significant differences between the histograms. Finally, we performed a cluster analysis using diffusion maps (Coifman et al., 2005) on the pooled trial data to predict the rigid and nonrigid layers as used in experimental conditions.

We exclude haze trials from further analyses due to the fact that between-subjects variability was high as described below. Unlike the other two layer conditions, haze trials did not indicate any grouping in one or two of the quadrants; instead, they were scattered. We analyzed the separation of pooled trials to decide whether there were any clustering in the data. To quantify the separation of pooled trials, we first determined the centroid per condition (water, haze, glass). Then we calculated L_1 distance (Minkowski distance, $p = 1$) of each condition's trial to its centroid. We present the sum of distances for all experiments in Table 1, as one can see, the highest sums and thus the highest separations are for the haze conditions in all experiments (values marked in bold). The sum of L_1 distances for haze trials were highest, suggesting that settings for this condition were scattered around all possible parameter values with no indication of
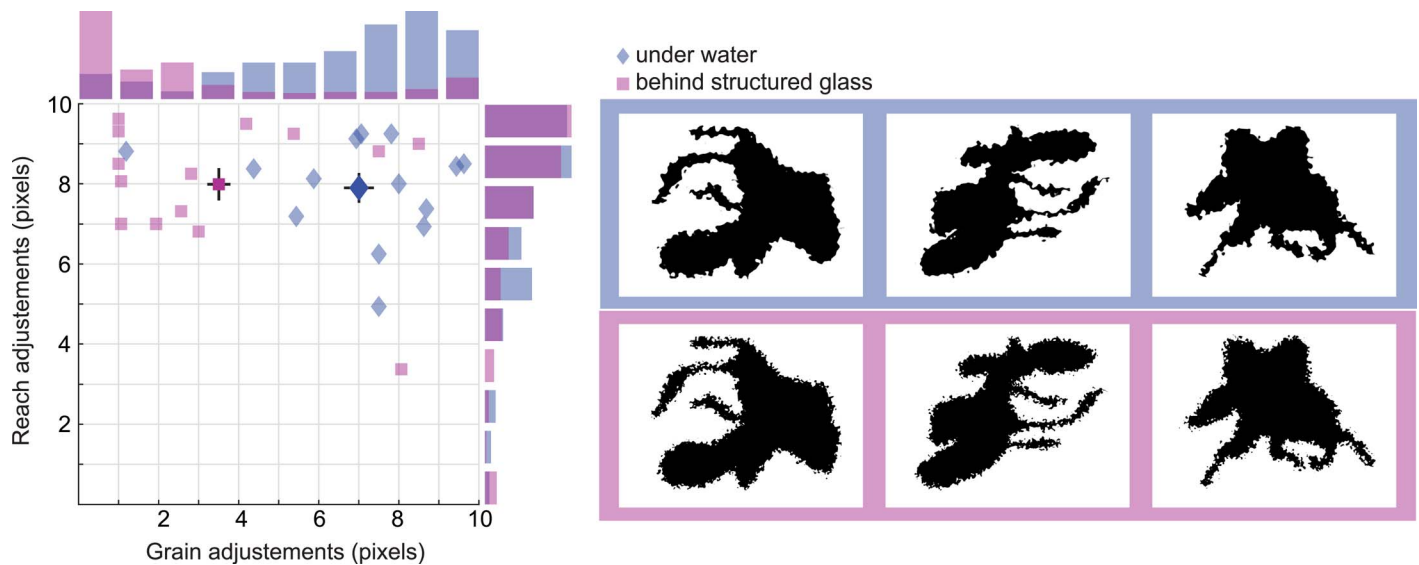
Figure 5. Individual and mean data from Experiment 1. Grain (horizontal axis) and reach (vertical axis) adjustments for each observer ($n = 14$) are plotted in separate icons for water (blue diamond) and glass (red square) trials. Each icon represents an observer's mean adjustments for all trials averaged over objects. Group means for each layer are also plotted with corresponding colors and larger icons ± standard error of the mean. On the right, example eidolons are shown for water (blue panel, top row) and glass (red panel, bottom row) that correspond to group mean settings. Note that the stimuli used in Experiment 1 has lower contrast, we use black and white values here to improve printed figures. The reach and grain settings of these examples are taken from the group means. Outside each axes, we plot normalized histograms for water (blue bars) and glass (red bars) trials pooled across participants. Histograms along the horizontal axis show data for grain adjustments, along the vertical axis we show histograms for reach adjustments.

clustering (see also Supplementary Appendix B, Supplementary Appendix Figure B1).

Observers

Fourteen naive observers participated in Experiment 1. All observers were from the Justus-Liebig-University Giessen, Germany (12 women, 2 men; mean age = 24.9 ± 3.5), with normal or corrected-to-normal vision. Observers gave written informed consent and were paid 6 Euros per hour for their participation.

Results

Figure 5 shows the group mean for grain and reach parameter adjustments (solid icons) overlaid with mean individual responses (transparent icons). For both grain and reach adjustments for water trials (blue icons), 11 out of 14 participants' settings are clustered in the upper right quadrant, suggesting that participants preferred large values for both parameters (μ_{grain} = 7.0, μ_{reach} = 7.9). Larger parameter adjustments generate wave-like image disarray as shown in the examples on the right panel of Figure 5 (right, blue panel). For glass trials, large reach adjustments were

frequently coupled with small grain values for all participants, except for three, (red icons, μ_{grain} = 3.5, μ_{reach} = 8.0). The group mean (large red icon) and corresponding "grainy" image deformations are also depicted in Figure 5 (right, red panel).

Another way of looking at the parameter adjustment results is to compare histograms for pooled trials. Discrepancy between water and glass trials are evident from separate distributions of grain adjustments (Figure 5, histograms along the horizontal axis). Glass trials present a narrow positively skewed histogram (red bars) and these are visibly separate from the negatively skewed wider distribution of water trials (blue bars). There is no separation for reach adjustments of water and glass (Figure 5, histograms along the vertical axis). Evidently, the separation between water and glass trials for reach settings is smaller than the separation between water and glass trials for grain settings: Wilcoxon test confirms that reach values for water and glass come from the same distribution ($Z = -1.5$, $p = 0.1$), where grain values for water and glass are different ($Z = 8.7$, $p < 10^{-17}$).

The stimuli in Experiment 1 differed in boundary shapes and sizes; despite these differences, observers preferred similar image disarray for all types of stimuli. We did not find an effect of shape on parameter adjustments, either for water or for glass trials

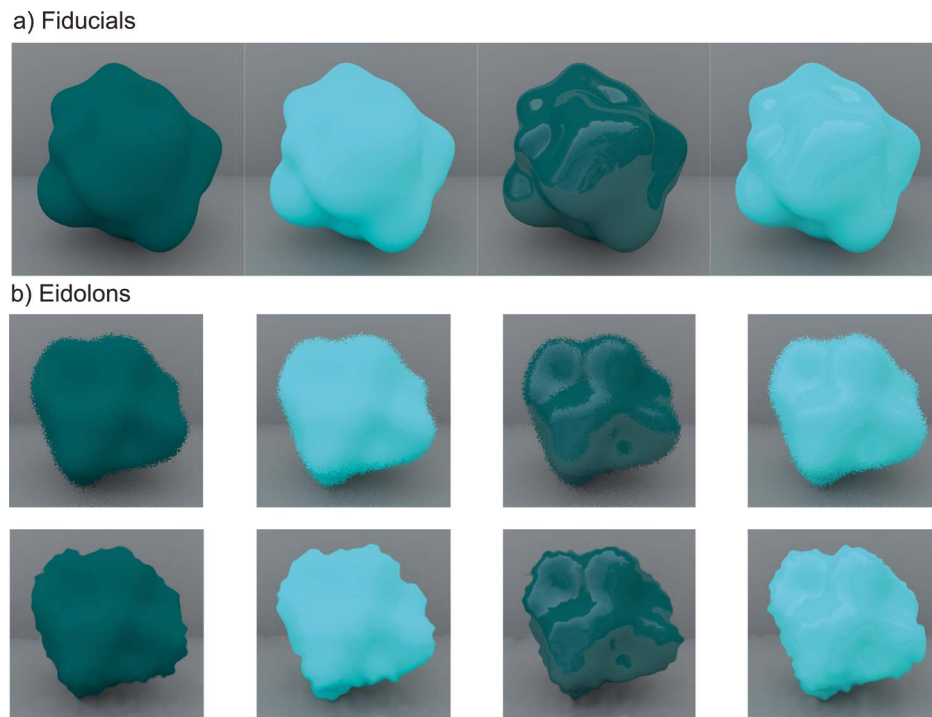a) Fiducials

b) Eidolons

Figure 6. Stimuli used in Experiment 2. (a) Screenshots of the 3D scenes were used as fiducial images. (b) Examples of coherent eidolons corresponding to the fiducial images in (a) with large reach, small grain settings (first row) and large reach, large grain settings (second row).

(Supplementary Appendix D, Supplementary Appendix Figure D1).

In half of the trials, screen coordinates were flipped to discourage participants from paying attention to screen locations instead of the image deformations. This did not make any effect on the parameter adjustments (see Supplementary Appendix A, Supplementary Appendix Figures A1 and A2).

Discussion

These results suggest that image distortions solely can yield a transparent layer percept, even in the absence of context. We also see the benefit of describing image distortions in two parameters: one parameter controls the locality and size of the group of pixels to be shifted whereas the other determines how much this group of pixels is shifted. The locality measure seems to make the difference according to the nature of the transparent layer. The difference in grain results suggests a clear separation between water and glass layers in terms of the locality of the disarray, whereas the similarity in reach settings suggests that the intensity of the image disarray is similar for both water and glass. We next investigate whether this trend transfers to more complex objects in Experiment 2.

Experiment 2: Glavens full

Motivation

After establishing the required image disarray for 2D shapes in Experiment 1, we wanted to explore whether having more information than a luminance defined boundary would result in similar parameter adjustments when perceiving a transparent medium. We introduced shape complexity to our stimulus set, in the sense that they convey 3D shape information with additional color and shading. Here a different group of participants ($n = 16$) adjusted reach and grain parameters for eidolons of 3D objects in order to make them look like they are seen under water and behind structured glass.

Stimuli

We used rendered 3D objects as fiducial images to create stimuli for Experiment 2 (Figure 6a). Fiducial images for the stimuli were images of 3D scenes rendered with spherical objects called *Glavens* (Phillips, Egan, & Perry, 2009). The surface of this globular object was rendered with Cycles Render Engine in Blender 2.73a, it was illuminated with a "campus" environment map

(Debevec, 2002). We used a mixed shader combining bidirectional scattering distribution (BSDF) nodes for diffuse (0% gloss) and glossy (10% gloss) components. The lightness value was 0.1 for dark objects, and 0.8 for light objects (roughness = 0, hue = 0.5, saturation = 1). To investigate the effects of (a) contrast on the object, we add dark glossy objects with high contrast on the surface due to highlights as compared to light glossy objects, and (b) contrast with the background, we add dark diffuse objects that present high contrast with the background due to lightness and light diffuse objects with lower contrast with the background. We had 27 viewpoints: 120° camera rotations on the x, y, and z axes, each participant viewed 4 out of these 27 viewpoints randomly. Images were 600×600 pixels in Experiment 2.

Again, the fiducial images were not shown in the experiments; instead, we used them to create eidolons that were used as stimuli (Koenderink et al., 2017). We produced a set of controlled image deformations to the fiducial images as in Experiment 1. In addition to coherent eidolons with sharper looking edges, we also used incoherent eidolons. This parameter controls the alignment of decomposed (and disarrayed) layers of the image. While a coherent disarray has aligned layers of coarse and fine structures, an incoherent disarray has shifted recomposition and as a result, the eidolons look fuzzier. This incoherence might give the impression of a less clear transparent medium, such as haze.

For each fiducial image we created a set of 100 eidolons (1–10 pixels for reach and grain) by using the eidolon factory (https://github.com/gestaltrevision/Eidolon). In Experiment 2, during the first test block, stimulus image disarray was coherent, whereas the second test block included incoherent disarray eidolons.

Observers

Sixteen different observers participated in Experiment 2. All observers were from the Justus-Liebig-University Giessen, Germany (11 women, 5 men; mean age = 24.8 ± 1.6), and had normal or corrected-to-normal vision. Observers gave written informed consent and they were paid 6 Euros per hour for their participation.

Procedure

As in Experiment 1, we used an adjustment method to see how observers varied image distortions for transparent layers of different nature.

In Experiment 2, participants completed 96 trials (2 reflectances $\times$ 2 lightnesses $\times$ 3 layers $\times$ 8 repetitions) per block (Figure 6). They first completed the coherent block, and then the incoherent block. All trials were interleaved in blocks in Experiment 2.

Analysis and data exclusion

The analysis procedure and exclusion of haze trials were as described in Experiment 1.

Results

Data from individual participants in Figure 7 (both blocks) lie strictly in the upper and lower right quadrants for water trials except from one participant (blue transparent icons). Likewise, for glass trials, individual data are mostly clustered in the upper left quadrant showing that participants preferred higher reach and lower grain values. These settings result in grainy image deformations as shown with example eidolons in Figure 7 (red panels on the right). We show the group mean for grain and reach parameter adjustments: In line with Experiment 1, for both coherence blocks, mean grain and reach adjustments for water trials (blue icons) are in the upper right quadrant. The majority of participants adjusted large values for the parameters in both coherent (Figure 7a; blue icons, $\mu_{grain} = 7.0$, $\mu_{reach} = 8.2$) and incoherent (Figure 7b; blue icons, $\mu_{grain} = 7.5$, $\mu_{reach} = 8.4$) blocks. For glass trials, similarly large reach adjustments were coupled with smaller grain values in both coherent (Figure 7a; red icons, $\mu_{grain} = 4.8$, $\mu_{reach} = 6.5$) and incoherent (Figure 7b; red icons, $\mu_{grain} = 4.6$, $\mu_{reach} = 7.3$) blocks.

Pooled trial histograms are very close to one another for reach adjustments, as can be seen from the normalized histograms along the vertical axes in Figure 7.

At a glance, for the coherent block (Figure 7a) the difference between water and glass histograms for reach values is smaller than for grain (along horizontal axis) adjustments. For the incoherent block (Figure 7b), again we observe a smaller difference for reach values (bars along vertical axis) than for grain (bars along horizontal axis) values. Indeed, the reach values for water and glass in the coherent block ($Z = 2.2$, $p < 0.04$) and the incoherent block ($Z = 0.6$, $p = 0.6$), both remain insignificant at a Bonferroni corrected level $\alpha < 0.0125$. Grain values for water and glass, on the other hand, are different in both blocks (coherent block: $Z = 16.4$, $p < 10^{-59}$, incoherent block: $Z = 17.8$, $p < 10^{-71}$). In accordance with Experiment 1, perceived water and glass layers were distinguished dominantly in the locality of disarrays but not the required distortion intensity.
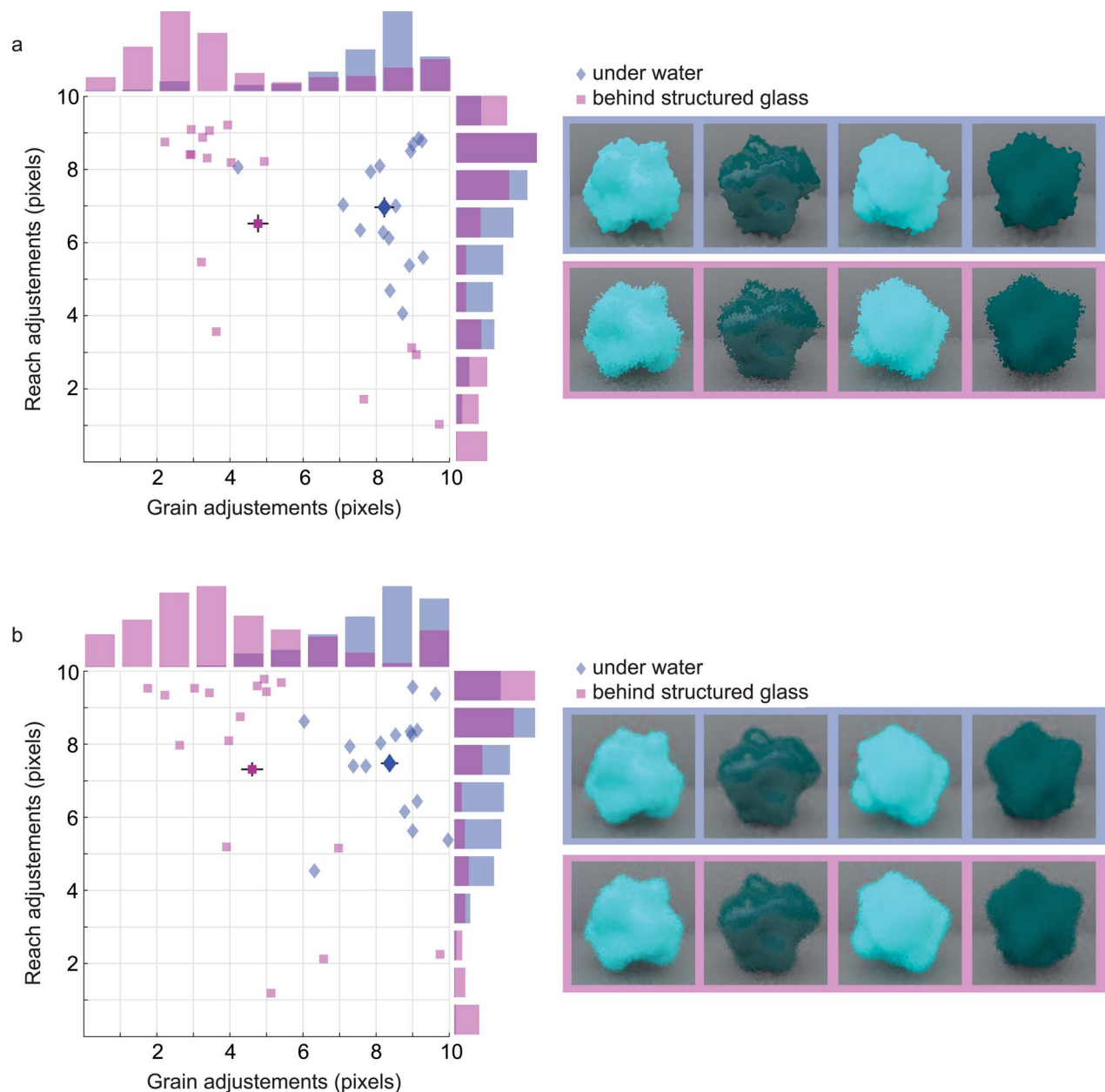
Figure 7. Individual and mean data from Experiment 2. (a) Coherent block: Grain (horizontal axis) and reach (vertical axis) adjustments for each observer (*n* = 16) are plotted in separate icons for water (blue) and glass (red) trials. Each icon represents an observer's mean adjustments for all trials averaged over objects. Group means for each layer are also plotted with corresponding colors and larger icons ± standard error of the mean. On the right, example eidolons are shown for water (blue panel, top row) and glass (red panel, bottom row) that depict the group mean settings for reach and grain. Normalized histograms for water (blue bars) and glass (red bars) trials pooled across participants are shown outside the axes. Histograms along the horizontal axis show data for grain adjustments and along the vertical axis, we show histograms for reach adjustments. (b) Results from incoherent block with same participants are shown in line with (a).

Coherent versus incoherent eidolons

When we contrast the coherence in water trials (Figure 7a vs. 7b, blue bars along the vertical axis), there are small but significant differences in reach values ($Z = -5.0$, $p < 10^{-6}$), where smaller difference grain values remain insignificant (same figure, blue bars along the horizontal axis, $Z = -2.3$, $p < 0.02$). Similarly, comparisons for coherent versus incoherent glass trials reveal a significant difference for reach values (same figure, vertical red bars, $Z = -11.0$, $p < 10^{-27}$); but not for grain values (horizontal red bars, $Z = 0.6$, $p = 0.6$). Overall, when we compare coherence of the eidolons, we only observe differences in the reach

settings: observers prefer higher reach values for incoherent eidolons in both water and glass trials.

Surface reflectance

In Experiment 2, we used four objects as stimuli (Figure 6a: GlossLight, GlossDark, DiffuseLight, DiffuseDark). When we compare the adjusted parameters for each object at a conservative significance level (Bonferroni correction: $\alpha < 0.05/24 = 0.0021$) we observe a single significant difference for the grain value comparisons for coherent water trials, between the DiffuseLight and GlossDark objects ($\Delta\mu G = 0.46$, $Z = 3.3$, $p < 10^{-4}$), and none for the reach values. Observers set smaller grain values for the dark object with highlights than for the light diffuse object. Detailed results for each object are shown separately in Supplementary Appendix D, Supplementary Appendix Figures D2 and D3.

Discussion

The addition of 3D surface properties to objects does not deviate from the parameter trends for transparent layers we find in Experiment 1: we observe large reach and grain values for perceived water layer, and large reach and small grain values for perceived glass layer. Moreover, in line with Experiment 1, we observe the difference between the two types of layers only in the grain parameter of image distortions.

We also note a difference between coherent and incoherent eidolons. These differences reveal that for both layers, observers set higher intensity image disarray for the incoherent stimuli, and the amount of difference is larger for the glass layer. Higher reach settings for the incoherent block can probably be related to the blurred edges of the incoherent stimuli. When there are no sharp edges, one can expect that observers set larger disarray to form above threshold image distortions for a transparent medium.

Finally, we also find an interesting difference in relation to the surface reflectance of objects: a dark glossy one behaves differently from a light diffuse one. At a glance, these two objects differ strikingly by the local contrast on their surfaces: the glossy dark object has highlights that show bright white patches on the dark surface color (hence, small grain settings suffice), whereas the diffuse light object has minimal local luminance contrast that makes it harder to perceive image disarray on its surface. We look further into the effects of 3D object properties on perceived transparency in Experiment 3.

Experiment 3: Glavens no boundary

Motivation

We found differences related to surface reflectance in Experiment 2. In Experiment 3, we reduced the surface related cues to 3D shape by occluding the object boundary. When there is no visible object boundary, observers rely on the reflectance to recover 3D shape curvature and we expect fine details such as surface highlights to play a crucial role here.

Stimuli and procedure

Similar to Experiment 2, stimulus set consisted of eidolons of glavens rendered with diffuse/glossy surface reflectance, and dark/light surface color. Essentially, the same eidolons as in Experiment 2 were shown behind a Gaussian aperture that occluded the object boundary (radius = 220 pixels). We repeated the same procedure as in the previous experiment with a different group of participants.

Observers

Ten different observers completed Experiment 3. All observers were from the Justus-Liebig-University Giessen, Germany (7 women, 3 men; mean age = 23.2 ± 1.8), and had normal or corrected-to-normal vision. Observers gave written informed consent and they were paid 6 Euros per hour for their participation.

Analysis and data exclusion

Analysis procedure and exclusion of haze trials were as described in Experiment 1.

Results

Overall, compared to Experiments 1 and 2, the data patterns for water and glass layers remained similar, with large reach and grain values for perceived water, and large reach values paired with small grain values for perceived glass. However, we observe lower reach adjustments paired with higher grain values for water trials and vice versa for glass trials. Detailed findings are reported below.

For both coherence blocks, mean grain and reach adjustments for water trials (Figure 8a and b, blue icons) are again in the upper right quadrant. Participants prefer large values for the parameters in both
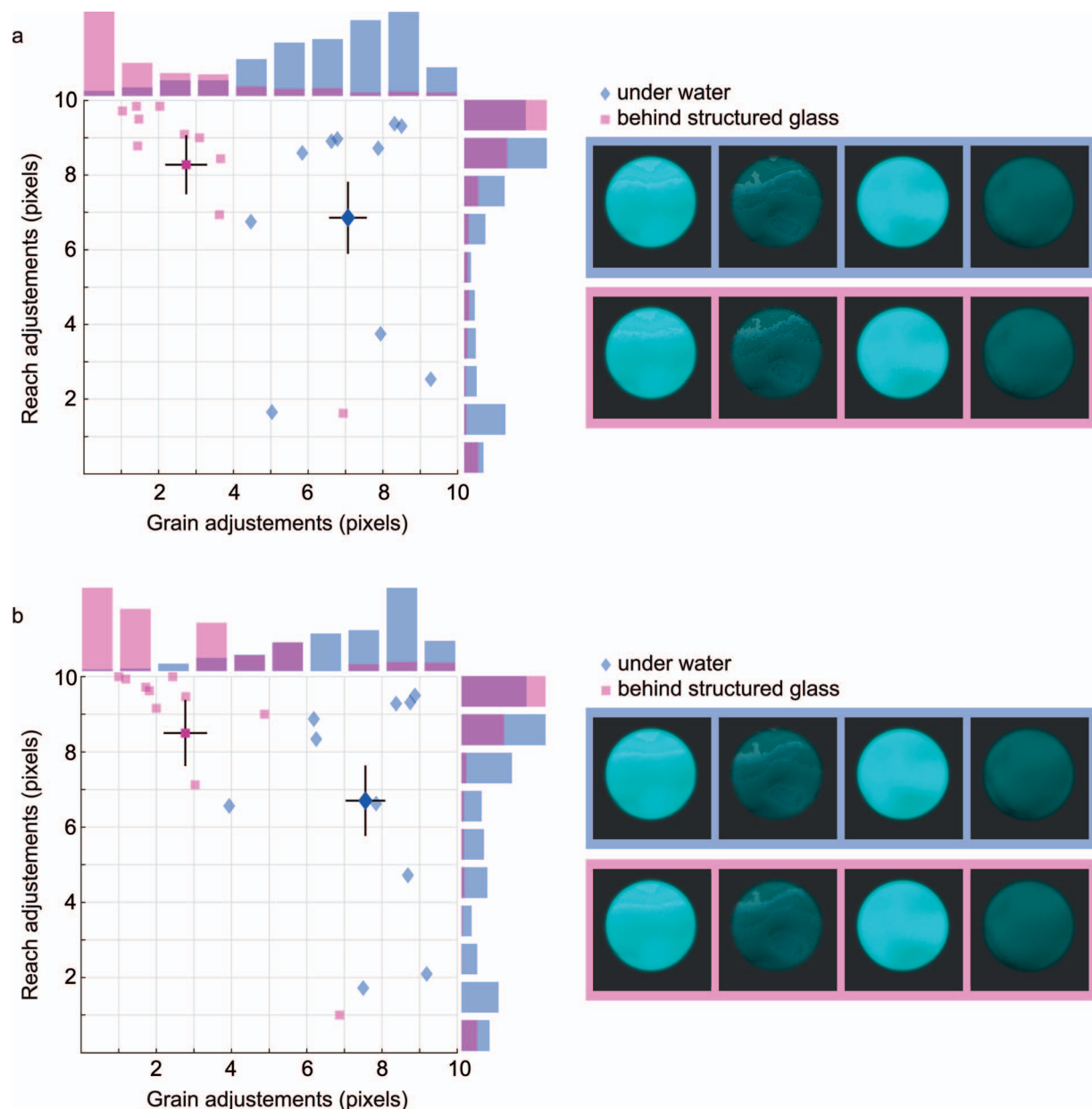
Figure 8. Individual and mean data from Experiment 3. (a) Coherent block: Grain (horizontal axis) and reach (vertical axis) adjustments for each observer ($n = 10$) are plotted in separate icons for water (blue) and glass (red) trials. Each icon represents an observers mean adjustments for all trials averaged over objects. Group means for each layer are also plotted with corresponding colors and larger icons ± standard error of the mean. On the right, example eidolons are shown for water (blue panel, top row) and glass (red panel, bottom row) trials' reach and grain means. Normalized histograms for water (blue bars) and glass (red bars) trials pooled across participants are shown outside the axes: horizontal histograms show data for grain adjustments, vertical histograms show reach adjustments. (b) Results from incoherent block with same participants are shown in line with (a).

coherent (Figure 8a, blue icon, $\mu_{grain} = 7.1$, $\mu_{reach} = 6.9$) and incoherent (Figure 8b, blue icon, $\mu_{grain} = 7.6$, $\mu_{reach} = 6.7$) blocks. For glass trials (Figure 8, red icons), similarly large reach adjustments were coupled with smaller grain values in both coherent (Figure 8a, red icon, $\mu_{grain} = 2.7$, $\mu_{reach} = 8.3$) and incoherent (Figure 8b, red icon, $\mu_{grain} = 2.8$, $\mu_{reach} = 8.5$) blocks. In comparison to the previous experiment, when the object boundary is occluded so more edges are missing,

the image distortions remain visible but very subtle (Figure 8, eidolons in the right panel).

The trial distributions are significantly separate in all of the conditions (Figure 8). For the coherent block (Figure 8a), as visible in the overlapping blue and red histograms, difference between water and glass histograms for reach values is smaller than for grain adjustments. For the incoherent block (Figure 8b), again we observe a smaller difference for reach values

than for grain values. The reach values for water and glass (histograms along the vertical axes) yield significant differences in distributions both in the coherent ($Z = -6.3$, $p < 10^{-9}$) and the incoherent block ($Z = -7.5$, $p < 10^{-13}$). Grain value distributions for water and glass are also different in both blocks (histograms along the horizontal axes, coherent: $Z = 14.2$, $p < 10^{-45}$, incoherent: $Z = 14.5$, $p < 10^{-47}$). As can be expected, observers consistently adjust larger grain settings for water trials in both blocks, which is in line with findings from our previous experiments. Surprisingly, this time we also find that observers set higher reach parameters for glass trials compared to water trials (in both blocks), a finding we have not observed in Experiment 1 or 2.

Coherent versus incoherent eidolons

Overall, coherent versus incoherent blocks reveal very small differences. In water trials, there is no significant difference in reach value distributions ($Z = 1.5$, $p = 0.1$), whereas grain values yield small but significant differences ($Z = -4.3$, $p < 10^{-4}$). Preferred grain settings are larger for the incoherent water stimuli when compared to coherent, but there is no distinction for coherence in reach adjustments of water trials. Comparison for coherent versus incoherent glass trials reveal small but significant differences for reach ($Z = -3.0$, $p < 0.003$), but differences for grain adjustments remain insignificant ($Z = -1.0$, $p = 0.3$). In contrast to our findings for water, the effect of coherence in glass trials are only evident in reach settings.

Surface reflectance

Similar to Experiment 2, pairwise comparisons reveal one object distinction at the conservative significance level of $\alpha < 0.0021$: the diffuse light object results in higher grain adjustments when compared to the diffuse dark object in coherent glass trials ($\Delta\mu G = 0.55$, $Z = 3.1$, $p < 0.002$). Detailed results for each object are shown separately in Supplementary Appendix D, Supplementary Appendix Figures D4 and D5.

Discussion

Pairings for reach and grain parameter settings remain in line with Experiments 1 and 2. We also find that layers are distinguished mostly by the grain parameter. Unlike the previous experiments, in this one we also observe distinctions in reach settings for the first time. One reason for this may be the missing object boundary information in the stimuli, the lack of the object outline might result in preferred image distortions to be more intense overall. The effect of eidolon coherence is trickier to interpret. Higher reach settings in the incoherent block may be related to the blurred outlines of the incoherent stimuli: when there are no sharp outlines, edge quality is lower, so one can expect that observers set a larger disarray to increase the intensity of image distortions and form a percept of a transparent medium. It is interesting to note that the increased magnitude in image distortions manifests itself in locality values for perceived water but in intensity values for perceived glass. One reason for this maybe because larger locality for perceived water are more natural as it can be related to larger ripples, whereas if the locality value of distortions are too large for perceived glass it might stop looking like glass and appear too diffuse.

Predicting transparent layer classes

Despite the variety in our stimuli throughout three experiments, the same trends in our findings prevail. By analyzing the two image disarray parameters, we see two separate trends for a rigid (glass) and a nonrigid (water) transparent layer. Next, we try to classify these parameter adjustments by using a diffusion map analysis (Coifman et al., 2005). A diffusion map analysis is done by a nonlinear dimensionality reduction algorithm that maps the elements of the dataset to a lower dimension Euclidean space via the eigenvalues of a Markov matrix. Then, by taking the diffusion coordinates into account, it performs k-means clustering on these data points. We predict two different clusters for a rigid and a nonrigid transparent layer, from the pooled data of (unlabelled) water and glass trials. An attempt to include haze trials and cluster data into three classes is included in Supplementary Appendix C.

In Figure 9, we present raw data of all trials from Experiments 1, 2, and 3, color coded for when participants adjusted grain and reach so that the layer looks like water (blue discs) and when the layer looks like glass (red discs). For each trial, the adjustments of the grain and the reach parameter are plotted on the horizontal and the vertical axes respectively. We then combined these settings and clustered them into two classes: the same figure shows superimposed circle icons for predicted water (blue circles) and glass (red circles) clusters. The predictions are strongly correlated with our measured trials for each experiment.

Figure 9a shows Experiment 1 results: out of 448 trials, 75.8% of water trials and 77.2% of glass trials fall into the predicted clusters, resulting in a significant correlation (Pearson's $r = 0.53$, $p < 10^{-33}$). For
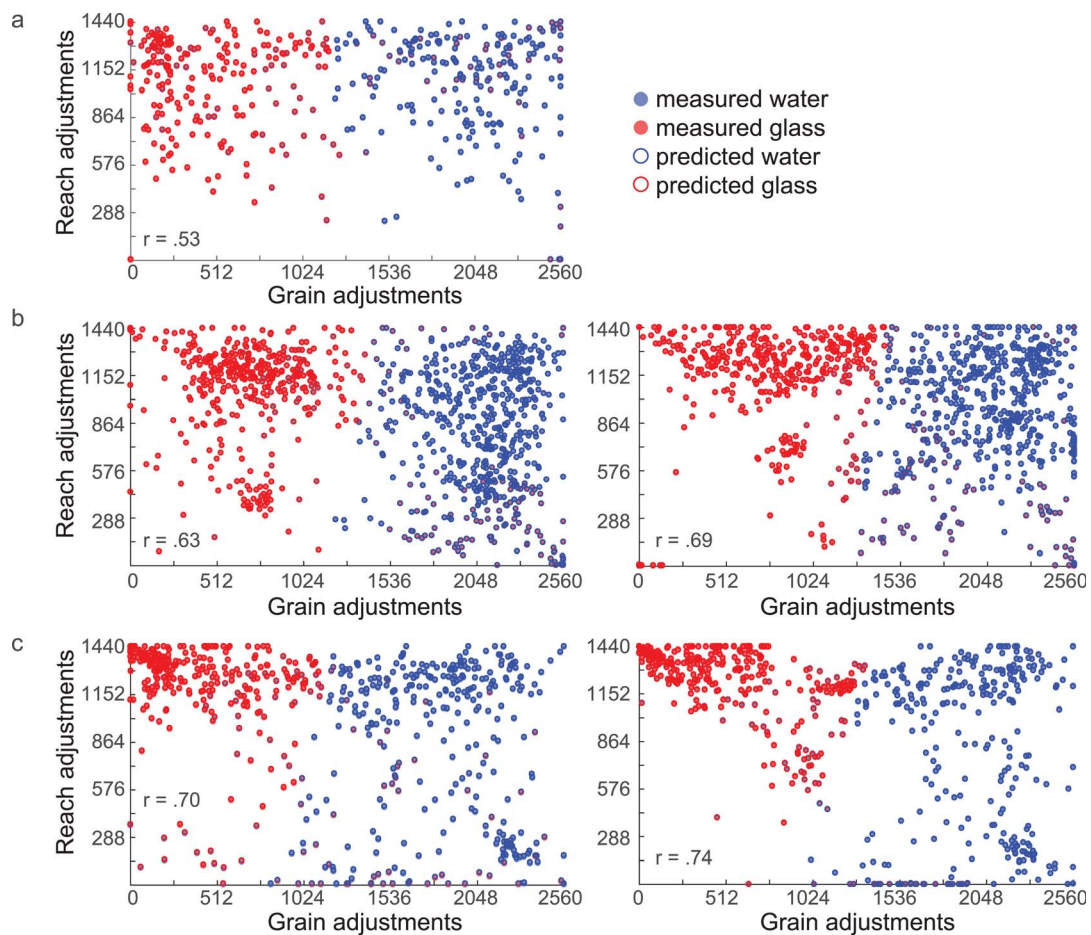
Figure 9. Predicted classes for water and glass from Experiments 1 through 3. (a) Grain (horizontal axis) and reach (vertical axis) adjustment data for all trials in Experiment 1. Filled icons depict trials where participants adjust for water (blue discs) and glass (red discs). Overlaying hollow icons show the predicted classifications for two clusters of water (blue circles) and glass (red circles). Correlation coefficients between measured and predicted classes are inscribed in each panel. (b) Same for trial data from Experiment 2, for the coherent (left panel) and the incoherent (right panel) blocks. (c) Same for trial data from Experiment 3, coherent (left panel) and incoherent (right panel) blocks.

Experiment 2 (Figure 9b), out of 1,024 trials, overlap is higher for water trials, and 90.8% in coherent block (left panel) and 91.9% in incoherent block (right panel) of them are predicted. Glass trials fall into the predicted cluster for 70.7% of the time in the coherent block, and 75.6% of the time in incoherent block. Overall correlation between experimental layers and predicted classes is significant in both coherent ($r = 0.63$, $p < 10^{-112}$) and incoherent ($r = 0.69$, $p < 10^{-142}$) blocks. Finally, in Experiment 3 (Figure 9c), out of 640 trials, 86.6% of water trials are predicted accurately in the coherent block (left panel) and 84.1% in the incoherent block (right panel). The predicted class for glass layer contains 83.1% of glass trials in the coherent block and 89.4% in the incoherent block. The correlations between predicted and actual classes are strong (coherent block, $r = 0.70$, $p < 10^{-93}$; incoherent block, $r = 0.74$, $p < 10^{-109}$).

Discussion

Transparent scenes often present rich visual details in our daily life, such as a swimming pool, a rainy windshield, or a shower screen. It is crucial to understand the nature of a transparent medium, for example, to distinguish rocks from the fish in a river. Here we first show that static image distortions alone can induce the percept of a transparent layer. Moreover, we parametrically manipulate these changes in images to identify different types of complex transparent layers.

We give a first account of how transparent layers in complex scenes can be explored in three adjustment experiments. High interobserver consistency in parameter adjustments confirm that the data trends in our findings are far from random. First, we show that when asked to imitate a water layer, participants consistently

pick large parameter values that result in large local radius, wave-like intense deformations. When asked to imitate a structured glass layer, the intensity of deformations remain similar but they consistently pick a smaller radius for local deformations which gives a grainy and diffuse look. For haze, adjustments between participants are less consistent (see below). This pattern transfers across a variety of stimuli from 2D luminance-defined shapes to 3D objects with more complex surface properties. Finally, when we cluster the parameter settings, our analyses reveal clearly separate classes for rigid (glass) and nonrigid (water) transparent layers that are overlapped by participants' choices. This suggests that the parameters are predictive in separating the rigidity of the perceived transparency.

Here, we report a simple way of describing perception of transparent layers in complex scenes. With only two parameters of local image deformations, namely locality and amplitude of the disarray, we can classify a rigid and a nonrigid transparent layer from participants' adjustments. When image deformations are larger (high grain values), the data indicate a nonrigid layer of water; participants perceive a rigid transparent layer (structured glass) when local deformations are smaller (low grain values). The intensity of deformations we measured are the same for both rigid and nonrigid transparent layers. Our experimental procedure allowed us to study the nature of local image disarray for perceived transparency, but it is not within the scope of this study to infer how strongly the observers perceived a transparent layer.

Recent evidence suggests that image deformations cause percepts of transparent layers in complex scenes including colored texture backgrounds (Fleming et al., 2011; Kawabe et al., 2015; Kawabe & Kogovšek, 2017). In Kawabe et al. (2015), the authors argue that image deformations are not enough to perceive transparent layers and that low-level motion signals are also required. Here, we provide several lines of proof that participants agree on the required image deformation parameters for convincing transparent layers, separately for a rigid and a nonrigid equivalence class, suggesting that static deformations are sufficient. One should note, however, that our interactive reach-grain adjustment task might have produced spurious dynamic information, which might have somewhat enhanced the transparency impression for observers.

Equivalence classes of transparent layers

Semantically, instances of a transparent layer that fall into the same equivalence class are as many as the richness of our lexicon: a swimming pool, a clear pond, a glass of water, or saline solution are just some examples that can be classified as water. This was also evident in the word lists from our brainstorming sessions. So, when testing for transparent layers, we avoided fully detailed verbal descriptions of water, haze, and structured glass. We also kept the 3D object geometry and scene properties, such as lighting or background, constant and as minimal as possible. This allowed us to emphasize image deformations in our stimuli, as small as a few pixels that were detected by the participants. We also intentionally avoided introducing background texture, because we aimed at experimenting how 2D and 3D objects are seen through a transparent material by manipulating object properties such as boundary and surface reflectance.

Background texture

With increasing focus in the literature on materials (Adelson 2001; for a review, see Fleming, 2017) and advances in computer graphics, it is now possible to explore layer decomposition with increased complexity. In naturalistic and complex scenes, it is often the case that the transparent medium has no intrinsic boundary visible (e.g., an outdoor scene on a foggy day), and the objects are immersed in the transparent medium. Transparency perception in complex scenes with composite material properties has found attention only in the last decade, most of the time using a blob detached from the background texture. Previous studies that focus on how we perceive thick transparent materials (Fleming et al., 2011; Schlüter & Faul, 2014, 2016) directly test the perceived thickness, shape, refractive index, and transparency of the transparent blob. Often they also report that the human visual system is sensitive to the properties related to the distorted texture background such as the density of texture and blob's distance from it. Kawabe et al. (2015) suggest that edge continuity and the shape of texture elements might cause a small difference in the spatiotemporal tuning of their dynamic image distortions (but see also Kawabe & Kogovšek, 2017). We believe that the background texture's spatial scale has an effect on the properties of the perceived transparent layer. To test this, we ran a pilot experiment with colored texture images from the McGill Database (Olmos & Kingdom, 2004). Figure 10 shows the results of a pilot experiment testing this idea (the complete set of results and analysis details can be found in Supplementary Appendix E, Supplementary Appendix Figures E1 and E2). It illustrates that participants tend to prefer smaller deformations for background textures with smaller elements, suggesting a noticeable relationship between texture element size and required image deformation. For example, textures 17 and 14 in Figure 10 contain both small pebbles and require exactly the same amount of disarray amplitude (reach
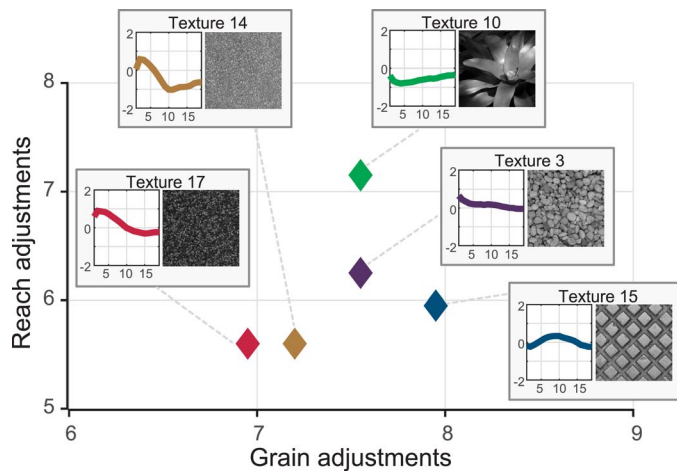
Figure 10. Reach and grain adjustments for water. Colored diamonds correspond to group means ($n = 10$) for five sample textures in the pilot texture experiment. Next to each symbol, the corresponding texture and the radial average of the textures' set mean-subtracted fast Fourier transforms (FFTs) (subtracting the set mean emphasizes differences in the amplitude spectra between textures) are shown. This figure illustrates that the frequency content of the texture affects observers' settings systematically. For example, settings for texture 14 (dark green) and 17 (orange) are quite close to each other and so are the corresponding set mean-subtracted radial averages of the textures' FFT. In contrast, the mean setting for texture 10 (light green) is further away from 14 and 17, and so is the shape of its FFT. Also see Supplementary Appendix E, Supplementary Appendix Figures E1 and E2.

values for dark green and orange icons), and are both far from texture 10, which contains large texture elements, (i.e., leaves). While it was not our aim here to search for a systematic relation between spatial content and image distortions required to obtain a specific apparent transparency class, these data suggest that there is a dependency on the spatial scale of the texture in images, such that smaller parameter settings are being made for textures with smaller element sizes.

Rigid versus nonrigid transparent layers

The visual system is very good at rapidly discriminating rigid and nonrigid motion (Jain & Zaidi, 2011), and we have a rich lexicon to describe these differences (e.g., turning vs. twisting, rotating vs. revolving). These 3D motion types manifest themselves as different deformation patterns in the optic flow; similarly, here we find that local image deformations differ for rigid and nonrigid transparent layers. Why is this the case?

Humans can recognize nonrigid objects (e.g., liquids from static images; Paulun, Kawabe, Nishida, & Fleming, 2015), or even make fine judgments, such as the viscosity of liquids from midlevel shape cues (van

Assen et al., 2018). It has been argued that image deformations can provide cues to perceived transparency, but weaker in the static case in comparison to dynamic image distortions (Kawabe et al., 2015; Kawabe, 2017).

More recently, the dependency between optical and mechanical properties of materials has been explored (Schmid & Doerschner, 2018), to provide evidence that mechanics of rigid and nonrigid materials can be inferred from static cues. Most of the time, optical properties in an image suffice for the way we expect materials to behave, except from chemical "miracle" examples (see Waitukaitis, & Jaeger, 2012; Grossman, 2013). When we think about rigidity in the scope of transparent layers, it is important to acknowledge that the visual system attributes motion information to nonrigid materials from local image deformations (e.g., image of an oozing liquid). Rigid and nonrigid materials behave very differently when exposed to external forces. Deformations of nonrigid materials under stress are continuous as a flowing river or gravity on raindrops, but rigid materials do not deform until a further limit; that is, they endure stress until a breaking point. For rigid transparent materials, local image deformations might mean structural patterns formed while nonrigid, as in ice or structured glass, or local deformations might suggest additional layers, such as sweat beads or raindrops on glass or plastic.

Our findings support that the visual system is very efficient to identify whether the image deformations are caused by rigid or nonrigid shape dynamics. We find that local image distortions, in general, are a useful way to study transparent materials as they systematically change with the materials' inferred motion or structure. Eidolons give more control over the local image disarray, making them better tools to study rigid transparent layers, Newtonian fluids (such as water or air), and non-Newtonian fluids that change their apparent viscosity under stress (such as honey or saliva).

Why are observer settings for haze all over the place?

Understanding and parametrically describing transparent layers in images are also useful for image processing algorithms to clean up noisy images, such as dehazing photographs while preserving color information (El Khoury, Thomas, & Mansouri, 2014). For this reason and for being a common atmospheric event, we included haze as one our transparency classes.

Participants were more consistent when adjusting image deformations for water and glass, compared to haze; hence, our classifications for the former were much better clustered than the haze parameterization.

One explanation for this might be because haze is semantically more complex than, for example, water by definition. Perhaps it is more ambiguous to visualize instances of haze compared to well-defined examples of water or glass. The physical cause of haze has to do with airlight: air becomes visible due to scattering of light, reducing the visibility of objects in the background. Mist, fog, snow, volcanic ash, widespread dust, sand, and haze are listed as classes of *obscuration* in international meteorological codes (World Meteorological Organization, 1995). Unlike mist, that is defined as dispersion of water, haze is the atmospheric phenomenon in which dry particles such as dust, sand, or volcanic ash are obscuring visibility, similar to fog. So, even though snow haze, caused by suspension of the snow particles in air, and dust haze, which can be caused by many things from ash to industrial gases, belong to the same class (haze), the two have rather different kinds of physical interaction. Therefore, the instances in an equivalence class of haze may have larger dissimilarities, whereas for water, one can think of many examples with similar refractive indices that fall into the same class. Thus, the ambiguity of haze might have caused a variety of image deformation settings in this study. The distortions when we see through a hazy transparent layer are sometimes blurry as in reduced contrast or intensity. Another explanation might be that if the observers wanted to mimic these types of distortions then our parameter range might not be suitable since we provided no control over the coherence parameter of eidolons in this study.

Conclusion

We focused on perceived transparent layers in complex scenes by way of image distortions. When objects are seen through water, glass, or haze, these transparent layers induce characteristic distortions in the image. In the case of water and glass, object contours deform whereas the edge quality and contrast remains unchanged. Haze, on the other hand, mostly reduces contrast without large affect on the shape of the contours. We used highly controlled image distortions to induce distinct impressions of these transparent layers. We found high interobserver consistency for image disarray settings, groups of observers readily agreed on perceived water and perceived glass layers. The high intraobserver consistency also suggested that our method captured parameter settings for perceived transparency reliably: since repetitions showed different objects and viewpoints, settings for perceived transparent layers cannot be reduced to pixelwise image properties. The settings were also robust in the sense that the patterns transferred across simple 2D shapes to

more complex 3D objects with, for example, glossy surfaces. Here we propose a novel method that, based on local image distortions, generates images that convey a vivid sense of transparency. Even though the images generated by this method may not appear as naturalistic as those produced by specialized 3D graphics software, our approach was very successful in identifying equivalence classes of water and glass, and our results were predictive to separate rigid from nonrigid transparent layers. Due to the richness in examples of complex transparent layers, we found it more efficient to inquire equivalence classes of, for example, water. Our results also suggest that complex transparency perception can effectively be studied with local image disarray parameters.

Keywords: transparency, layer decomposition, eidolon factory, image disarray, image distortion, surface reflectance

References

Adelson, E. H. (2001). On seeing stuff: The perception of materials by humans and machines. In: B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of SPIE: Human Vision and Electronic Imaging VI: Vol. 4299* (pp. 1–12).

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, 26(4), 419–453.

Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., & Zucker, S. W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion

maps. *Proceedings of the National Academy of Sciences, USA, 102*(21), 7426–7431.

Debevec, P. (2002). Image-based lighting. *IEEE Computer Graphics and Applications, 22*(2), 26–34.

Dövencioğlu, D. N., Welchman, A. E., & Schofield, A. J. (2013). Perceptual learning of second order cues for layer decomposition. *Vision Research, 77*, 1–9.

El Khoury, J., Thomas, J.-B., & Mansouri, A. (2014). Does dehazing model preserve color information? In *Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS) 2014* (pp. 606–613). Marrakesh, Morocco: *IEEE.*

Fleming, R. W. (2017). Material perception. *Annual Review of Vision Science, 3*(1), 365–388. pMID: 28697677. https://doi.org/10.1146/annurev-vision-102016-061429.

Fleming, R. W., Jäkel, F., & Maloney, L. T. (2011). Visual perception of thick transparent materials. *Psychological Science, 22*(6), 812–820.

Grossman, L. (2013). Movies of miracle gloop show it shatters like glass. *NewScientist, 218*(2913), 12.

Jain, A., & Zaidi, Q. (2011). Discerning nonrigid 3d shapes from motion cues. *Proceedings of the National Academy of Sciences, USA, 108*(4), 1663–1668.

Kawabe, T. (2017). What property of the contour of a deforming region biases percepts toward liquid? *Frontiers in Psychology, 8*: 1014.

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

Kawabe, T., Maruya, K., & Nishida, S. (2015). Perceptual transparency from image deformation. *Proceedings of the National Academy of Sciences, USA, 112*(33), E4620–E4627.

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.

Kingdom, F. A. (2011). Lightness, brightness, and transparency: A quarter century of new ideas, captivating demonstrations, and unrelenting controversy. *Vision Research, 51*(7), 652–673.

Kleiner, M., Brainard, D., & Pelli, D., Ingling, R., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception, 36*, 1–16, Supplement 14.

Koenderink, J., Valsecchi, M., van Doorn, A., Wagemans, J., & Gegenfurtner, K. (2017). Eidolons: Novel stimuli for vision research. *Journal of Vision, 17*(2):7, 1–36, https://doi.org/10.1167/17.2.7. [PubMed] [Article]

Koenderink, J. J., & van Doorn, A. J. (2001). Shading in the case of translucent objects. In *Human Vision and Electronic Imaging VI, Vol. 4299* (pp. 312–321). International Society for Optics and Photonics.

Latecki, L. J., Lakamper, R., & Eckhardt, T. (2000). Shape descriptors for non-rigid shapes with a single closed contour. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2000, Vol. 1* (pp. 424–429). Hilton Head Island, SC: IEEE.

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics, 13*(1), 59–66.

Morgenstern, Y., Schmidt, F., & Fleming, R. (2017). Effects of shape transformations on perceived similarity. *Journal of Vision, 17*(10): 1383, https://doi.org/10.1167/17.10.1383. [Abstract]

Olmos, A., & Kingdom, F. A. (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception, 33*(12), 1463–1473.

Paulun, V. C., Kawabe, T., Nishida, S., & Fleming, R. W. (2015). Seeing liquids from static snapshots. *Vision Research, 115*, 163–174.

Phillips, F., Egan, E., & Perry, B. (2009). Perceptual equivalence between vision and touch is complexity dependent. *Acta Psychologica, 132*(3), 259–266.

Rowling, J. K. (2014). *Harry Potter and the deathly hallows. Vol. 7.* London: Bloomsbury Publishing.

Schlüter, N., & Faul, F. (2014). Are optical distortions used as a cue for material properties of thick transparent objects? *Journal of Vision, 14*(14):2, 1–14, https://doi.org/10.1167/14.14.2. [PubMed] [Article]

Schlüter, N., & Faul, F. (2016). Matching the material of transparent objects: The role of background distortions. *i-Perception, 7*(5): 2041669516669616, https://doi10.1177/2041669516669616.

Schmid, A. C., & Doerschner, K. (2018). Shatter and splatter: The contribution of mechanical and optical properties to the perception of soft and hard breaking materials. *Journal of Vision, 18*(1):14, 1–32, https://doi.org/10.1167/18.1.14. [PubMed] [Article]

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research, 39*(16), 2697–2716.

Schofield, A. J., Hesse, G., Rock, P. B., & Georgeson, M. A. (2006). Local luminance amplitude modulates the interpretation of shape-from-shading in textured surfaces. *Vision Research, 46*(20), 3462–3482.

Schofield, A. J., Rock, P. B., Sun, P., Jiang, X., & Georgeson, M. A. (2010). What is second-order vision for? Discriminating illumination versus material changes. *Journal of Vision*, *10*(9):2, 1–18, https://doi.org/10.1167/10.9.2. [PubMed] [Article]

Tamura, H., Higashi, H., & Nakauchi, S. (2018). Dynamic visual cues for differentiating mirror and glass. *Scientific Reports*, *8*(1): 8403.

van Assen, J. J. R., Barla, P., & Fleming, R. W. (2018). Visual features in the perception of liquids. *Current Biology*, *28*(3), 452–458.

Waitukaitis, S. R., & Jaeger, H. M. (2012, July 11). Impact-activated solidification of dense suspensions via dynamic jamming fronts. *Nature*, *487*(7406), 205–209.

World Meteorological Organization. (1995). *Manual on codes: International codes. Vol. 1.* Geneva, Switzerland: Secretariat of the World Meteorological Organization.

Zhou, Y., & Baker, C. (1996). Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat. *Journal of Neurophysiology*, *75*(3), 1038–1050.