https://doi.org/10.2352/EI.2022.34.15.COLOR-222

Exploring the role of caustics in translucency perception — An eye tracking approach

Davit Gigilashvili, Aditya Suneel Sole, Shaikat Deb Nath, Marius Pedersen;
Department of Computer Science, Norwegian University of Science and Technology; Gjøvik, Norway;

Abstract

Translucency is an important appearance attribute. The caustic patterns that are cast by translucent objects onto another surface encapsulate information about subsurface light transport properties of a material. A previous study demonstrated that objects placed on a white surface are considered more translucent by human observers than identical objects placed on a black surface. The authors propose the lack of caustics as a potential explanation for these discrepancies — since a perfectly black surface, unlike its white counterpart, does not permit observation of the caustics. We hypothesize that caustics are salient image cues to perceived translucency, and they attract the visual attention of the human observers when assessing translucency of an object. To test this hypothesis, we replicated the experiment reported in the previous study, but in addition to collecting the observer responses, we also conducted eye tracking during the experiment. This study has revealed that although gaze fixation patterns differ between white and black floor images, the objects' body still attract most of the fixations, while caustics might be a cue of only secondary importance.

Introduction

Translucency is one of the essential appearance attributes that plays an important role in how objects and materials appear, but our understanding of how the human visual system perceives translucency remains limited [1]. Two-dimensional caustic patterns are formed when *"light is reflected (or refracted) at one or more specular surfaces, focused into ray bundles of a certain structure, and then received as patterns of light on a diffuse surface."* [2] The structure of the light that is transmitted through a translucent object and projected onto another surface encapsulates information about this object (e.g. refer to Fig. 1). Previous studies have reported that human observers explicitly use caustic patterns or lack thereof for assessing translucency of a material [3, 4]. Gigilashvili *et al.* [5] conducted psychophysical experiments to explore the correlation between caustics and perceived translucency. They used physically-based rendering to generate images of translucent objects, and asked observers to assess their translucency. The same object was shown in two different environments — placed on a white floor with clearly visible caustics below the objects, or placed on a perfectly black floor that absorbed all light projected onto it, making it impossible to observe the caustics. The authors found that the objects placed on a white floor were judged significantly more translucent than their black-floor counterparts.

The results reported in [5] could be an indication that caustics are significant cues to translucency. However, the data did not permit drawing strong conclusions, as changing the floor color changes the overall luminance distribution in the entire scene, affecting other image cues as well (refer to Fig. 2). Hence, it was not possible to determine whether all variation in the observer responses could be solely attributed to the absence of caustics. In this work, we hypothesize that caustics are salient image cues to translucency that human observers attend to when making translucency assessments. The primary objective of the study is to test this hypothesis. Identifying other potentially informative image regions for translucency and patterns in observers' behavior are additional objectives. We replicated the experiment reported in [5] with an addition of tracking observers' gaze. Eye tracking has been previously proposed as a promising way to identify the regions in the image structure that are cues to translucency [1, 4]. However, to the best of our knowledge, this is the first work conducting actual eye tracking for translucency perception research.

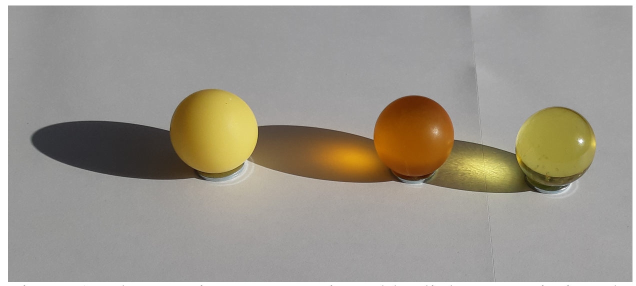

Figure 1: The caustic pattern projected by light-transmissive objects onto other surfaces encapsulates information about the materials these objects are made of. While opaque objects cast dark shadows, colorful caustic patterns are visible below transparent and translucent objects. Reproduced from [6]. © 2020, Society for Imaging Science and Technology

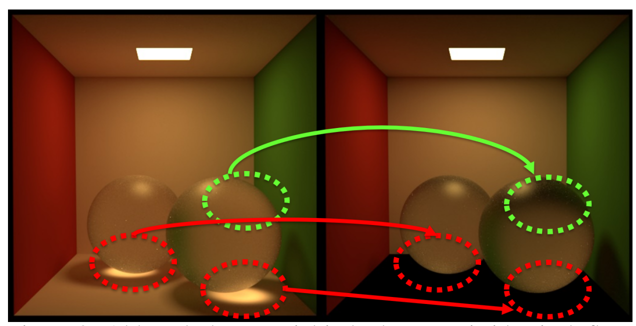

Figure 2: Although the material in both scenes is identical, floor color changes its appearance. The black floor not only removes caustics (marked red), but also considerably affects overall luminance distribution on the objects (marked green), which the reported perceptual difference might also be attributed to.

The article is organized as follows: in the next section, we describe the methodology and the experimental setup. In the subsequent sections, we analyze whether the quantitative results of the experimental task have been reproduced from the previous experiment, followed by the presentation of the eye tracking results. Afterward, we discuss the findings and outline future directions.

Methodology

Stimuli and Experimental Protocol

We replicated the psychophysical experiment reported in [5], using the same visual stimuli and experimental protocol. Physically-based renderings of six different translucent materials that come in five different shapes have been shown in a Cornell box with white and black floors (refer to Fig. 3-4). The task in the category judgement psychometric scaling experiment has been to assess the translucency of the material on a 1-to-6 scale. As reported previously, scaling translucency is a challenging task, as no universal standard exists how *"more translucent"* should be interpreted [1, 7]. Therefore, in this experiment, we explicitly imposed a direction to the scale, defining *less translucent* as *"closer to opacity"*. Besides, judging absolute magnitude of translucency is also highly subject to individual interpretation. To ensure relative judgment, as in the previous study, a reference of the two extremes in the dataset has been also displayed throughout the entire experiment. Refer to [5] for the detailed experimental protocol.

Observers

22 observers, including 3 co-authors of this paper, have participated in the experiment – 17 males and 5 females with a normal or corrected-to-normal vision and an average age of 31.41 years. The original study reported in [5] had 50 observers. Only one observer has participated in both experiments.

Observation Conditions

Although the study was again hosted at the QuickEval [8] platform, unlike the original experiment, which was conducted online, this experiment took place under controlled laboratory conditions. The setup is illustrated in Fig. 5. The stimuli were displayed on a color-calibrated 24.1 inch EIZO ColorEdge CG246 monitor. Whitepoint was calibrated to 6500K color temperature and 80cd/m^2. The display was the only light source in the experimental room. Chin rest located 60 cm away from the monitor was used to minimize head movements and to ensure that observers' pupils were continuously detected by an eye tracking device. The displayed images occupied approximately 14.4° of the field-of-view both vertically and horizontally.

Eye tracking

A stationary Gazepoint GP3 HD eye tracker with 60Hz sampling rate was used to record the movement of both pupils and the points of gaze. Before conducting each experiment, we made sure the pupil was detected and conducted 9-point camera calibration in Gazepoint Control software. Although the manufacturer claims visual angle accuracy to be in the range of 0.5-1.0 degree, we achieved the maximum accuracy of 2 degrees in the center, with higher inaccuracy in the top left and top right corners. This was deemed acceptable as the test and reference images are located in the center of the display.

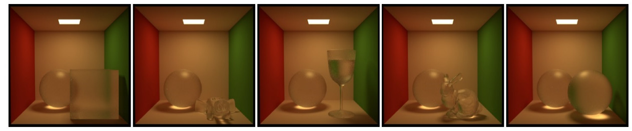

Figure 3: The materials were presented in five different shapes: cube, elephant, wine glass, Stanford Bunny, and sphere. A sphere made of the same material was also present in all scenes. Reproduced from [5]. © 2020, Society for Imaging Science and Technology

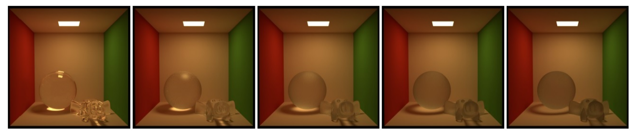

Figure 4: Translucency was modulated with surface scattering. In addition to perfectly smooth objects, the ones with the room-mean-square slope of microfacets equal to 0.2, 0.4, 0.6, and 0.8 have been used (shown in the elephant shape, from left to right, respectively). Reproduced from [5]. © 2020, Society for Imaging Science and Technology

Analysis

Firstly, mean observer scores were calculated to assess the reproducibility of the results reported in [5]. Secondly, gaze and pupil information were analyzed. We studied how the number of gaze fixations varied across different test images and how the fixations on the test and reference images relate. Moreover, we constructed gaze maps to determine to what extent do the observers attend to the caustic patterns and what are other image regions of potential interest. The maps were constructed in the following way: gaze point coordinates were extracted for each fixation and the respective point in the map was assigned the value of gaze duration. The cumulative map of all 22 observers was constructed for each test image. Considering visual acuity and eye tracker accuracy limitations, we do not gaze at particular pixels – but the surrounding areas. Therefore, a Gaussian filter with kernel size equivalent to the 2° of the visual field and standard deviation of 5 was applied to the raw gaze map. Finally, the center of gravity [9] was found for each gaze map.

Figure 5: The setup used for the experiment. A stationary eye tracking device is installed below the monitor. The distance between the chin rest and the monitor is 60cm.

Table 1: Pearson and Spearman linear correlation coefficients between mean observer scores of the previous and new experiments. It is apparent that the results are generally highly correlated. The strongest correlation is observed for Cubic and Wine Glass shapes (green), while it's weakest for the Elephant (red). Elephant is the most complex shape, which seemingly leaves more room for subjective interpretation.

	Pearson	Spearman
All data	0.98	0.97
White floors	0.98	0.97
Black floors	0.99	0.97
Sphere	0.99	0.97
Bunny	0.98	0.99
Glass	0.99	0.99
Elephant	0.97	0.93
Cube	0.99	0.99

Results

In this section, we first present the quantitative results of the observer responses and compare with the results of the previous experiment. Afterward, we report how the number of fixations was changing throughout the experiment. Finally, we analyze the most attended parts of the image with gaze maps.

Reproducibility of the previous results

The mean observer scores for the previous and new experiments are illustrated in Fig. 6. The results show that although the 95% confidence intervals are broader in the recent study, the overall trend remains the same – the material is usually considered less translucent when it is shown on a black floor. The larger confidence intervals can be explained with the fact that only 22 observers participated in the recent study, while original study had 50 observers. The confidence interval (CI) is calculated as shown in Eq. 1 and takes the number of observations into consideration, more observers leading to smaller confidence intervals.

$$CI = 1.96 \times \frac{\sigma}{\sqrt{N}}, \quad (1)$$

where σ is the standard deviation, and N is the number of observations. Table 1 shows that the results of the two experiments are highly correlated.

Number of fixations over time

Analysis of the total number of fixations for each test image reveals an interesting behavioral trend. Fig. 7 illustrates that the number of fixations on the first trial is considerably higher than on the subsequent trials – slowly decreasing as the experiment progresses. Furthermore, for each trial, the portion of the fixations on the reference image is initially large in relation to the test image and then decreases asymptotically (refer to Fig. 8). This means that the observers scrupulously inspect the interface and the images on the first trial to get familiar with the task and the dataset. In the beginning of the experiment, observers need to attend to the reference image (which remains the same throughout the experiment) to get a sense of the relative scale. However, they eventually remember how the extremes of the scale look and inspect mostly the test image. In other words, observers get trained, and all subsequent tasks take less and less effort. A similar phenomenon was previously reported and included in the qualitative model of material appearance as an *adaptation* category in [3].

Gaze map

Considering that the drop-down menu the observers are using to select the answer and the *Next* button they use for navigation between the trials are naturally attracting the gaze, they are irrelevant for the hypothesis in question and bias the overall gaze map. Therefore, only the horizontal section displaying the test and reference images is considered. The centers of gravity of the gaze maps are shown in Fig. 9, while the maps for all white floor and all black floor images are illustrated in Fig. 10. The intensity of the original frame is decreased by 50%, and the normalized gaze map is overlaid for visualization's sake – white pixels in the gaze map correspond to the areas with most duration-weighted fixations, while transparent areas correspond to the areas with no fixations. The figure shows the cumulative gaze map, where fixations of all white floor and black floor images are summarized, respectively.

It is apparent that the cumulative gaze maps differ between the two floors. We conducted statistical sign tests on the two cumulative gaze maps. We applied Bonferroni correction to account for multiple testing, and found that the two maps are significantly different ($p < 0.01$). Moreover, we conducted a similar test for all 30 objects between the maps for its white floor and black floor versions, and found that white floor and black floor gaze maps are significantly different for 29 out of 30 identical object pairs. The only exception is the roughest Elephant object.

The gaze maps show that fixation points are slightly higher in the black floor images. While some fixations are on the floors in the caustic area as hypothesized, they amount a small portion of all fixations. Most of the time observers' gaze was directed to the left spherical object. The center of gravity (see Fig. 9) is slightly higher for the while floor, but falls in both cases on the left spherical object.

It is worth mentioning that the cumulative gaze map is overlayed on a spherical test image for visualization's sake. At first glance, it seems counter-intuitive that many fixations fall on the wall above the right spherical object. However, these fixations in the cumulative map can be attributed to shapes other than a sphere, such as a wine glass, which extends further up horizontally.

The previous study intriguingly proposed that even when the fully absorbing black makes it impossible to observe the caustics on the floor, observers might be using the remaining caustics on the wall. A small portion of fixations are noticable on the green wall when the floor is black, but it is orders of magnitude fewer than the ones on the object itself.

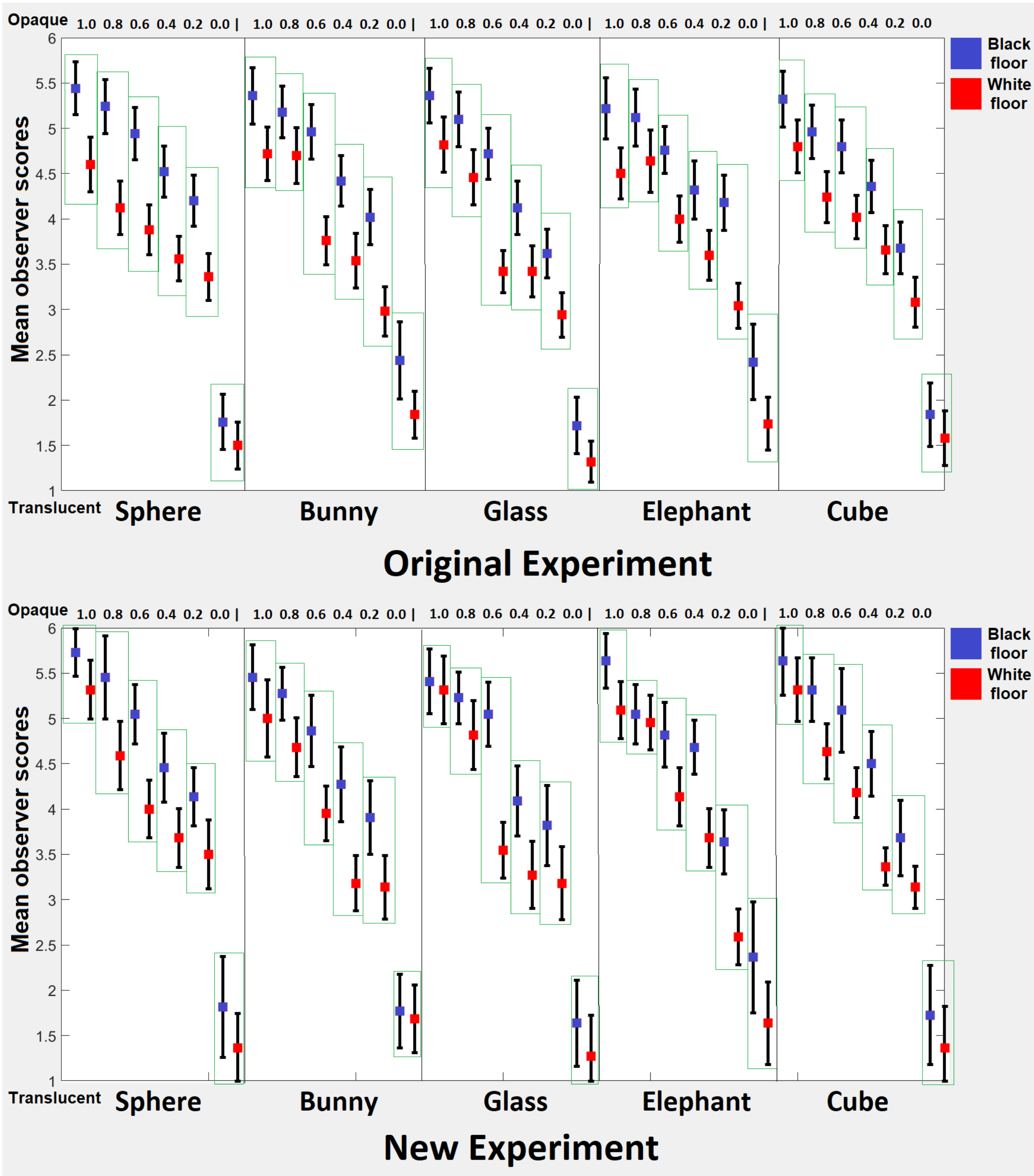

Figure 6: The mean observer scores for the original experiment reported in [5] and the new experiment. Blue squares correspond to black floor, while red squares mark the objects on a white floor. Whiskers extend to 95% confidence intervals. The confidence intervals are clipped whenever they go below 1 and above 6 (observers could choose categories between 1-6). The numbers on the top correspond to surface roughness *alpha*. The pairs of an identical material shown on two different floors are framed with a green rectangle for visualization's sake. Although a smaller number of participants resulted in larger confidence intervals, the overall trend remains the same in both experiments – black floor usually leading to a higher score.

Discussion

The observed gaze patterns do not provide a conclusive answer to the question raised in the previous publication whether the variation in observers' responses *"can be attributed to caustics only"*. The observers indeed look at the floor, including the caustic areas. They inspect the floor in black floor images as well, although to a smaller extent. This might be explained with their expectation learned from white floor images to look for the caustics. However, against our hypothesis this is not the part of the image with most fixations. Most fixations are usually attended to the left spherical object, which makes it unlikely that caustics are the primary cue in these images. However, considering the inaccuracy of 2° of the visual field, the spatial resolution of the image does not permit us identify exactly which part of the spherical ob-

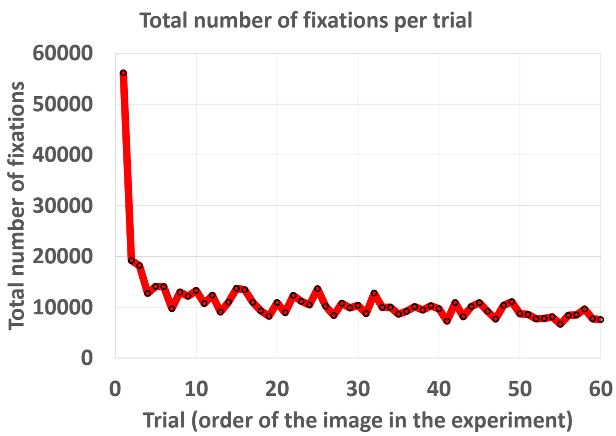

Figure 7: Total number of fixations for each trial. The numbers on the horizontal axis correspond to the order of the image shown during the experiment. The order has been identical for all observers. It is apparent that observers inspect the first trial scrupulously, while subsequent trials gradually need less and less gaze fixations.

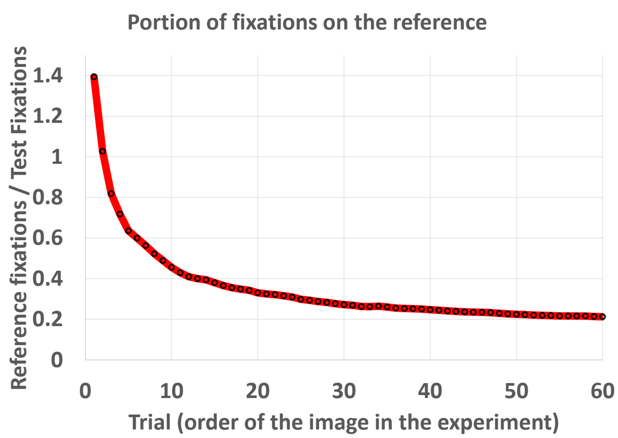

Figure 8: Horizontal axis corresponds to the order of the test images in the experiment. Vertical axis corresponds to the portion of the total number of fixations on the images that is attended to the reference image. Observers seemingly remember the reference and pay less and less attention to it as the experiment progresses.

ject is considered most attention-worthy and how this relates to the state-of-the-art knowledge on translucency perception. Considering that the materials in the dataset vary in surface roughness only, while subsurface scattering is zero, we cannot even rule out that observers are inadvertently assessing gloss and sharpness of the specular reflections, instead of translucency. Interestingly, the data fully supports the proposal made in the previous publication that absence of cross-shape differences can be attributed to the presence of the spherical object in all images. It turns out that observers are simplifying the task and instead of locating various complex shapes on the translucency-opacity scale, they rely on the shape, which remains invariant and which is used for definition of the scale range in the reference image. This is especially apparent for complex shapes, such as a rough elephant, which is the only object were gaze maps didn't significantly differ between the floor colors – as observers are simply looking to the same spherical object in both cases.

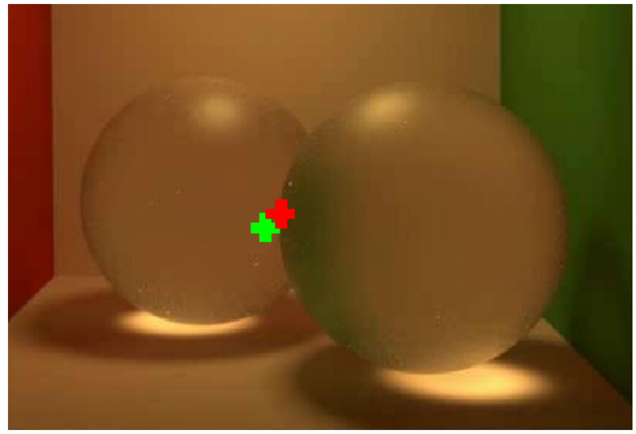

Figure 9: Centers of gravity of the cumulative gaze maps for all white floor (green) and black floor (red) images. The point is slightly further away from the floor when the floor is black, but in most cases, both points fall on the left spherical object.

Translucency is a complex psychovisual phenomenon, which proposedly involves weighted combination of multiple image cues. One potential hypothesis why caustics in these images attracted less attention in comparison with the physical objects in real-life experiments is the fact that they are overly simplistic and contain less information about objects' color and surface coarseness in comparison to the real objects visualized in Fig. 1. Another striking difference between Fig. 1 and synthetic Cornell box images that could proposedly explain the decrease in caustics' role in comparison with previous studies [3, 6] is the illumination angle – in Fig. 1 the grazing angle illumination ensures that caustics are detailed and spread over a large surface, while the overhead lighting in the Cornell box produces just small and homogeneous caustics [10]. It is surprising how quickly observers master solving the task. This, however, might not be the case if the trials are less predictable and environments as well as materials are more diverse, complex and natural – making us conclude that the role of caustics in the complex mechanism of translucency perception requires further investigation with higher spatial resolution, and more complex scenes and materials.

Conclusion and Future Work

We revealed a very interesting pattern in observers' behavior – they tend to simplify the task by judging the invariant spherical objects instead of variable shapes, which explains why no cross-shape differences have been detected. Moreover, observers put more effort in the beginning. However, they get familiar with the scale and need less fixations per trial as the experiment progresses. On the other hand, the study was not able to provide a decisive answer to the question, whether the variation in observer responses can be attributed to caustics only. It seems more likely that caustics act as a secondary cue, while primary cues to translucency are based on the luminance distribution on the object proper. Future work should address materials with subsurface scattering, use images with larger spatial resolution and more complex environment. This might shed more light to the question exactly which image regions, both on and outside the objects, are informative for translucency assessment and whether these regions vary across shapes and different environments.

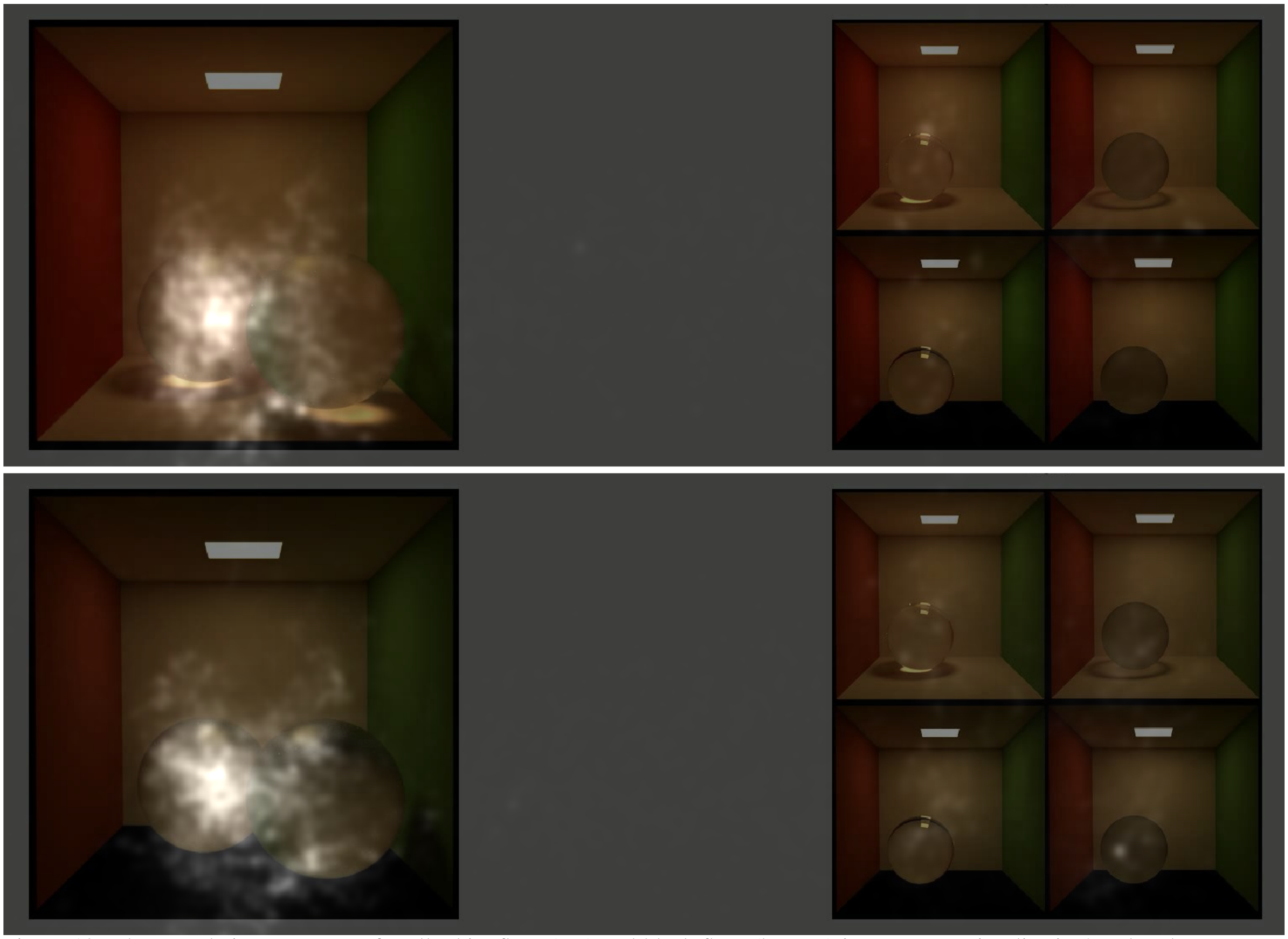

Figure 10: The cumulative gaze maps for all white floor (top) and black floor (bottom) images. For visualization's sake, the maps are overlaid on the respective images of a spherical object. The intensity of the background image is decreased by 50%, while the gaze map is normalized – the points with the largest number of fixations corresponding to white and the ones with no fixations remaining transparent. The map shows that while caustics attract part of the observers' gaze, most of the time, observers look at the object itself. Interestingly, observers mostly look at the left object, which is always spherical and not at the right object, which varies in shape. The two maps are significantly different at the 99% confidence level.

References

[1] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, Translucency perception: A review, Journal of Vision, 21, 8:4, pp. 1-41 (2021).

[2] Michael Wand and Wolfgang Straßer, Real-time caustics, Computer Graphics Forum, 22, 3, pp. 611-620, Wiley Online Library (2003).

[3] Davit Gigilashvili, Jean-Baptiste Thomas, Marius Pedersen, and Jon Yngve Hardeberg, On the appearance of objects and materials: Qualitative analysis of experimental observations, Journal of the International Colour Association (JAIC), 27, pp. 26–55 (2021).

[4] Davit Gigilashvili, On the Appearance of Translucent Objects: Perception and Assessment by Human Observers, PhD thesis, Norwegian University of Science Technology, Gjøvik, Norway, 2021.

[5] Davit Gigilashvili, Lucas Dubouchet, Marius Pedersen, and Jon Yngve Hardeberg, Caustics and translucency perception, Proceedings of Material Appearance 2020, IS&T International Symposium on Electronic Imaging, Society for Imaging Science and Technology, pp. 033:1–033:6. (2020).

[6] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, Behavioral investigation of visual appearance assessment, Proceedings of the 26^{th} Color and Imaging Conference. Society for Imaging Science and Technology, pp. 294–299. (2018).

[7] Davit Gigilashvili, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Marius Pedersen, On the nature of perceptual translucency, Proceedings of the 8^{th} Annual Workshop on Material Appearance Modeling (MAM2020), Eurographics Digital Library, pp. 17–20. (2020).

[8] Khai Van Ngo, Jehans Jr. Storvik, Christopher André Dokkeberg, Ivar Farup, and Marius Pedersen, Quickeval: a web application for psychometric scaling experiments, Proceedings of Image Quality and System Performance XII, International Society for Optics and Photonics, 9396, pp. 93960O–1–93960O–13. (2015).

[9] Jered Wells, Center of Mass, https://www.mathworks.com/matlabcentral/fileexchange/41675-center-of-mass. Retrieved: January 29, 2022.

[10] Oral comments by Prof. Mathieu Hébert during the discussion session on January 19^{th}, 2022, in the Electronic Imaging Symposium 2022.