TRANSPARENT OBJECTS: INFLUENCE OF SHAPE AND COLOR ON DEPTH PERCEPTION

Ar-pha Pisanpeeti, Éric Dinet

Université de Lyon, UJM Saint-Etienne
CNRS, UMR 5516, Laboratoire Hubert Curien, F-42000 Saint-Etienne, France

ABSTRACT

Recovering depth information from a single still image is an important problem in computer vision. However, the problem is difficult and challenging because it has an infinite number of solutions. To address this issue, humans use numerous visual cues to infer depth. Much progress has been made towards an understanding of the visual mechanisms involved in 3D perception. Such an understanding provides relevant knowledge to design efficient approaches for computer vision. While there is much prior work on opaque objects, there has been relatively little in relation with transparency. In this paper we investigate depth estimation from single still images containing nonplanar and real transparent objects. We focused our study on two visual features: shape and color. A database of stimuli was created to carry out a psychophysical experiment with 42 naïve observers.

Index terms— Perceptual transparency, depth perception, shape, color, psychophysical experiment

I. INTRODUCTION

Depth perception relates to the human visual ability to perceive the world in three dimensions (3D) and to evaluate object distances and sizes [1]. The main theory of depth perception is the Depth Cue Theory which states that the human visual system (HVS) combines different features, called depth cues, to reconstruct the third dimension from the 2D images projected on the surface of the retinas [2]. Another theory, the Ecological Theory, states that the visual environment and the current visual task play a significant role in the integration of distinctive objects into a coherent visual scene and in depth perception [3-5].

Both theories agree that the HVS uses different sources of information about the 3D nature of the observer's surroundings. Even if the perceptual process of depth sensing is still an open issue today [6-8], sources of information are generally divided into three types: pictorial, oculomotor and stereo depth cues [9]. Pictorial depth cues correspond to all image features that can feed the 3D reconstruction (e.g. linear perspective) [10]. Oculomotor depth cues refer to the feedback to the brain from the eye muscles. Stereo depth cues rely on the dissimilarities between the images on the left and on the right retina. All these depth cues are combined by the HVS to create the sense of 3D.

In literature, works on the perception of depth have focused on opaque objects, only few studies relates to transparent objects. However, this makes intuitive sense to think that the perception of depth and the perception of transparency are linked and that they interact each other [11]. Many researchers investigated perceptual transparency using images containing only simple flat filters. In 1970, Metelli proposed the episcotister model, which is the first quantitative model of the photometric constraints for perceptual transparency [12, 13]. He pointed out the role of: i) unity of transparent regions, ii) continuity of boundaries, iii) adequate stratification. Later, Gerbino *et al*. and Zanforlin *et al*. showed that photometric / chromatic conditions are necessary but not sufficient for perceiving transparency [14, 15]. They suggested that both luminance values and chromatic constraints have to be considered rather than only the reflectance or lightness values. Besides that, other studies explored the role and the influence of intersecting edges known as "X junctions" in the processes giving rise to perception of transparency [16-19].

During the last decade, many works attempted to refine theories of perceptual transparency. For example, Ripamonti *et al*. reviewed the necessary conditions for the perception of transparency and described the spatiochromatic constraints for achromatic and chromatic transparent displays [20]. They related the perception of transparency to the general question of how observers are able to correctly recognize the color of a surface when such a color is altered. More recently, Anderson argued that the transparency is only referred when there is visual evidence of a perturbation in local luminance and/or contrast along continuous contours [21].

The main difficulty with all these studies lies in the fact that human beings do not manipulate or interact with planar flat filters in their everyday life but with 3D transparent and translucent objects (e.g. glasses). Kraft *et al*. claimed that perceptual transparency in real scenes involves many additional factors such as surface specularities or the degree of spatial uniformity of surfaces [22]. This means that all models derived from studies only based on flat filters cannot clearly predict the perceptual transparency and translucency of real 3D objects. To explore such an issue, Nagai *et al.* studied the spatial regions related to translucency perception judgments [23]. They computed images of 3D objects on the basis of a psychophysical reverse-correlation method. Their results suggest that the rms contrast from whole stimuli is not important and that the local luminance could not be used as a critical cue because it induces only small effects.

The different works on perceptual transparency have pointed out the complexity of the topic especially because it involves from low-level to high-level image features. As for the perception of depth, perception of transparency certainly uses different sources of information from the observer's surroundings [8, 24-26]. What and how all the cues are combined and how they interact is still poorly understood.

In the work presented in this paper, we investigated the influence of perceptual transparency on the perception of depth from 2D images. Singh *et al.* pointed out "the dangers of identifying perceptual theories with the inverse of physical models" [19]. Then, we used real 3D transparent objects and captured their image under controlled lighting conditions to create a database of stimuli. With this database we carried out a psychophysical experiment with 42 naïve observers. We focused on two features of transparent objects: shape and color. If the HVS makes use of such features derived from perceptual transparency, then we should expect judgements of depth to be affected by changes in the transparency characteristics put in the psychophysical experiment.

II. Experiment

Stimuli — We selected 15 real transparent objects to create a database of stimuli. In order to investigate the influence of transparent shape on depth perception, two families of 7 glasses were chosen, one with a straight shape and the other one with a curved shape, as shown in Fig. 1. Similarly, to investigate the influence of transparent color in depth perception, the glasses cover a wide range of different hues.

Fig. 1. Real transparent objects used to create a database of stimuli

Stimuli consisted of monocular images of the selected nonplanar objects captured in controlled conditions. A transparent bell dome (left part of Fig. 1) was placed at a fixed position in front of a diffuse white background in a light booth uniformly illuminated by a D65 light source. As shown in Fig. 2, each glass mentioned above was aligned with the bell dome in 3 distinct positions varying from "in front" to "inside" and "behind". A Canon EOS D550 color camera was used to capture the images for the 3 positions of the 14 glasses with the same working distance, the same lens parameters, the same exposure time, and the same white balance previously processed on the white background.

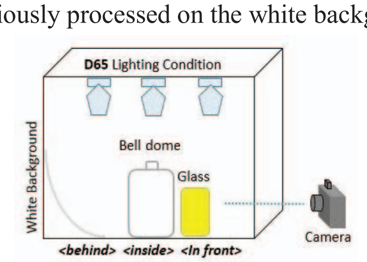

Fig. 2. Representation of the image capture setup. Each glass was shot for 3 positions: in front, inside and behind the transparent bell dome

In order to be able to determine the respective influence of shape and color features on depth perception, the database was completed by adding a cropped version of each original captured image. Such cropped images, that we refer to partial-context stimuli, have color information as a dominating transparency cue. Monocular cues in relation with the object shape such as relative size or linear perspective are almost absent in partial-context stimuli while they are strong in the full-context stimuli (original images), as shown in Fig. 3.

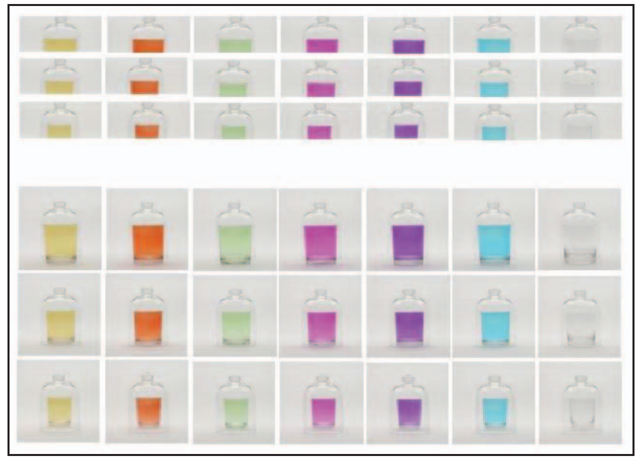

Fig. 3. Some stimuli of the database used to carry out the psychophysical experiment. The top part corresponds to the partial-context stimuli and the bottom part to the full-context stimuli.

The full database contains 84 stimuli corresponding to the combination of the 3 positions of each 14 transparent object with the 2 viewing contexts (full and partial).

Participants — Forty two naive observers (24 males and 18 females) with no object's prior knowledge participated in the psychophysical experiment. All subjects had normal or corrected-to-normal visual acuity and no colorblindness. Participants were all volunteers.

Equipment — Stimuli were displayed on a 24'' calibrated EIZO CG241W LCD color monitor (1920 × 1080 pixels) at a viewing distance of 55 cm. The experiment was carried out in a dark room. The surrounding luminance was kept at 10 $cd.m^{-2}$ with a luminance of 100 $cd.m^{-2}$ for the background of images. Matlab and Psychotoolbox were used to generate the stimulus sequences and to record the observer's responses collected through a three-switch experiment keyboard. Subjects viewed the screen with both eyes and the height of the observer's chair was adjusted to align the eye with the center of the monitor.

Procedure — On each trial, participants were presented with one stimulus centered on the screen and asked to answer the question: "Is the small object located in-front, inside or behind the big object?". For statistical consistency, subjects viewed each stimulus three times. Observers were given unlimited time to respond to each trial and provided their answer by pressing one of the three active keys of the keyboard. No further information about the objects was specified. Before collecting the data, participants were given a few practice runs with stimuli showing quite different transparent objects than those used to create the database presented above.

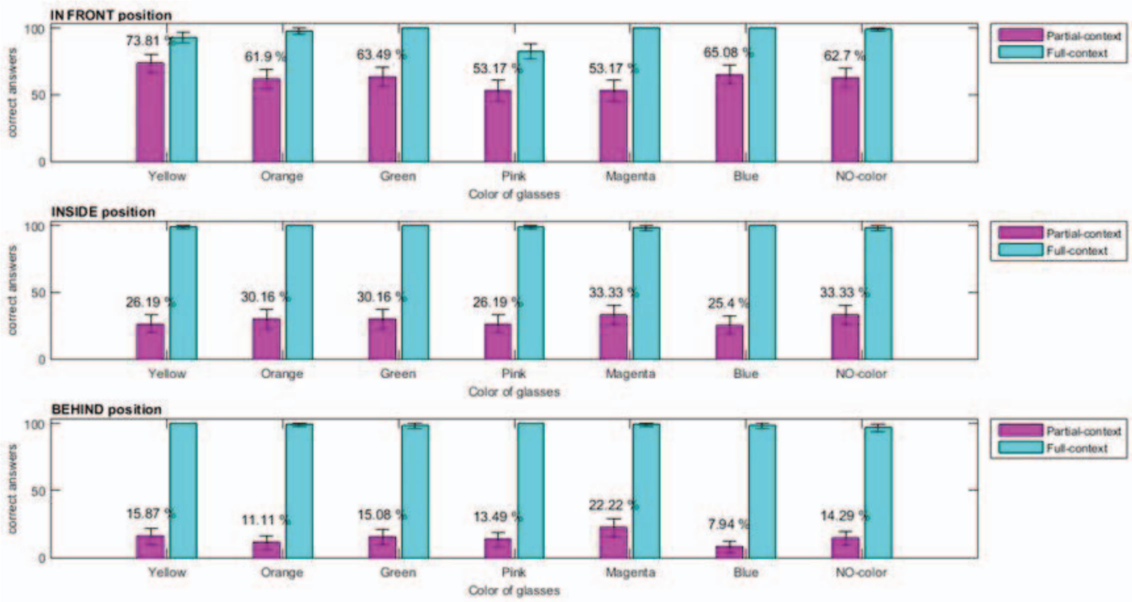

Fig. 4. Percentages of correct answers given by the 42 observers when judging the position of the 7 glasses with a straight shape relatively to the bell dome ("in front", "inside" and "behind"). Blue bars correspond to the full-context stimuli and pink bars to the partial-context stimuli

The psychophysical experiment was divided into two sessions. During the first session, only the 42 partial-context stimuli were randomly displayed across 126 trials. Observers were not able to predict the next stimulus. The first session lasted approximately 20 min. At the end of this first part of the experiment, subjects were allowed to take a break of 5 to 10 min. Then the second session started with the 42 full-context stimuli randomly displayed across 126 trials. The second session lasted approximately 10 min.

III. Results

A first global analysis of all the data collected during the psychophysical experiment shows that observers responded exactly the same at least 2 out 3 trials for each stimulus. Then, we can assume that all participants' responses are consistent to be used in our study.

Influence of shape on depth perception — Fig. 4 plots the percentages of correct responses obtained from the stimuli presenting the 7 glasses having a straight shape. For all configurations, it is clear that full-context stimuli (blue bars) lead to better scores than partial-context stimuli (pink bars). Such a general observation is confirmed by ANOVA and t-test ($p < 0.01$). Similar results were obtained with stimuli presenting the 7 glasses having a curved shape.

When the glass is located behind the bell dome, the partial-context is not sufficient to the observers for an efficient estimation of the correct relative position of the two transparent objects. In the full-context stimuli, observers can easily detect common monocular cues such as relative size or linear perspective. These monocular cues are not specific to transparency but they are strong evidence of depth for the expertise of the HVS. For the behind position presented through the partial-context stimuli, all the percentages of correct answer are under the chance level (see bottom graphs shown in Fig. 4). This suggests that observers tried to compensate the absence of monocular cues but that they failed in efficiently analyzing the visual information. Only on the basis of data in relation with transparency, observers misinterpreted the scenes. A short interview of each participant at the end of the psychophysical experiment revealed that most of them tried to find distortions or light reflections. They explained they rapidly identified the curved shape of the bell dome and they hypothesized that such a shape should alter and transform the structure of all objects located behind. Observers assumed that the straight edges of the glasses would be curved if the bell dome was located in front. Such an assumption induces an overestimation of the "in front" position because the bell dome did not introduce any distortion.

Subjective information collected during the individual interviews are confirmed by a statistical analysis of the data. In the collection of transparent objects used to construct the database of stimuli, three glasses with a straight shape have their respective colors identical to three glasses with a curved shape. Such a characteristic was measured and checked with a Minolta CS100 spectroradiometer. The two glasses of a pair differ in shape but not in color and the three pairs offer a panel of three distinct colors (blue, yellow and orange). For the analysis we also used the two neutral transparent glasses which have the same transmittance. A variance analysis and a t-test were conducted on the percentages of correct responses obtained by the 42 participants when estimating the position of the glasses relatively to the bell dome from partial-context stimuli. The statistical conclusions agreed the subjective information given by the observers. The "in front" judgement is significantly affected by the shape of the glasses but not by the color ($p < 0.01$). The differences reported for the two other configurations ("inside" and "behind") were not significant. By assuming that the bell dome tends to bend the straight lines viewed through it, participants underestimated the “in front” position for the glasses having a curved shape. With no additional monocular cue, the estimation of the relative position of real transparent objects is biased by the interpretation induced by their perceived shapes.

Influence of color on depth perception — To have an overview for each of the 14 real transparent glasses, confusion matrices were derived from the responses collected for the 42 configurations of the partial-context stimuli. Table 1 presents for example confusion matrices obtained with the yellow glasses. Columns correspond to the ground truth (GT) and rows to observers’ judgement (Obs). The highlighted diagonal cells represent the correct answers and the values to percentages of responses.

Straight YELLOW			
GT/Obs	In-front	Inside	Behind
In-front	73.81	21.43	4.76
Inside	69.05	26.19	4.76
Behind	33.33	50.79	15.87

Curved YELLOW			
GT/Obs	In-front	Inside	Behind
In-front	46.83	38.10	15.08
Inside	46.83	37.30	15.87
Behind	13.49	72.22	14.29

Tab. 1 Confusion matrices for partial-context stimuli

When the real yellow glass having a straight shape is placed in front of the bell dome, approximately 74% of responses are correct while when it is inside the bell dome, approximately 69% of responses are confused with “in front”. When this yellow transparent object is placed behind the bell dome, only about 16% of responses are correct and approximately 51% of responses are confused with “inside”. Similar trends can be observed for all other colored glasses but not for the two clear (non-colored) glasses. The straight clear glass tends to be always perceived in front of the bell dome. This is consistent with results derived from the statistical study focused on the shape of transparent objects and discussed above. The straight edges of the glass are interpreted as seen directly and not through the bell dome supposed to introduce geometrical distortions. Such an assumption can also explain the confusion matrix of the curved clear glass. Observers had difficulties in accurately estimating the position of this glass especially when it was placed inside or behind the bell dome. They interpreted the curved edges of the glass as a possible alteration induced by the bell dome and they found no clear additional visual information to easily select the right location. However, we can suppose that some luminance changes due to differences in light absorption may contribute to the final judgement given by the observers especially to decide between “inside” and “behind”.

Data presented in the confusion matrices were collected during the first session of the psychophysical experiment i.e. from the responses obtained when participants viewed only partial-context stimuli. At this stage of the experiment, observers had no prior knowledge about the transparent objects (shape, size, etc.). They were only instructed to answer the question: “Is the small object located in-front, inside or behind the big object?”. In such a context, confusion matrices show that all colored transparent objects were perceived closer than they really were. The local variations in the shape of the edges of the colored area of the stimuli seem not to have a preponderant influence when estimating the relative position of transparent objects. Results suggest that transparent color information affects the depth perception when the global shape of objects is not clearly identified by observers.

In the same context and in the absence of color information, contours induced by local changes in light absorption are the only low-level features on which observers can base their interpretation of the scene. Contours are in relation with shape and they are analyzed as such by the HVS. Data collected from the experiment suggest that color information may occult shape information in the context of transparency when shape is not intelligible i.e. when it does not play a significant role in the integration of distinctive and coherent objects, and when it does not induce monocular cues.

IV. Discussion

The work presented in this paper relates to the problem of depth perception involving perceptual transparency. The goal was to investigate the influence of shape and color features on the human visual ability to recover depth information from a single still image.

We created a database of stimuli by using real nonplanar transparent objects of two types: glasses with a straight shape and glasses with a curved shape. By selecting a sufficient number of different colors and simple shapes we aimed to focus our study only on the influence of targeted visual features. Stimuli were obtained by capturing images of the transparent objects arranged in different 3D configurations. We provided minimal information about stimuli and simple instructions to participants in order to reduce risks of bias. The visual task was only to estimate the relative position of two transparent objects. The psychophysical experiment started with partial-context stimuli for which observers were not able to see the global shape of transparent objects. The next stage of the experiment consisted in presenting full-context stimuli displaying the global shape of scene elements.

The collected data showed consistency within and across the 42 observers who volunteered the experiment. The partial-context stimuli revealed that color of transparent object could alter depth perception by giving an illusion of nearness. This only happens when shape features are not prominent and do not convey significant information for the integration of coherent objects. This is consistent with the findings of Guibal and Dresp who pointed out that “color is not an independent depth cue, but is strongly influenced by luminance contrast and stimulus geometry” [27].

The full-context stimuli revealed that object shapes were the dominant visual information in recovering depth. Familiar objects (transparent glasses and bell dome) were easily identified by observers and the visual analysis of their respective shapes bring additional geometrical cues. These cues were used to interpret the scene and its depth. The past experience and the knowledge of observers is a key element in the process. Participants hypothesize about the reciprocal influence of the transparent objects and the shape distortion it could induce. Hypotheses can be valid and help the observer to correctly recover depth information or they can be incorrect and induce a wrong perception of depth.

REFERENCES

[1] I.P Howard, “Perceiving in depth, volume 1: basic mechanisms,” Oxford University Press, 2012.

[2] G. Westheimer, “Location and line orientation as distinguishable primitives in spatial vision,” Proc. Roy. Soc. London, Series B, Biological sciences, vol. 263, no. 1369, pp. 503-508, 1996.

[3] J. Gibson, “The ecological approach to visual perception,” London: Lawrence Erlbaum Associates, 1986.

[4] M. Wexler and J.J.A. van Boxtel, “Depth perception by the active observer,” Trends in Cognitive Sciences, vol. 9, pp. 431-438, 2005.

[5] P. Sterzer and G. Rees, “Perceived size matters,” Nature Neurosciences, vol. 9, pp. 302-304, 2006.

[6] W.H. Ittleson and F.P. Kilpatrick, “Experiments in perception,” Sci. Am., vol. 185, pp. 50-55, 1951.

[7] M.S. Livingstone and D.H. Hubel, “Psychophysical evidence for separate channels for the perception of form, color, movement and depth,” J. Neuroscience, vol. 7, pp. 3416-3468, 1987.

[8] J. Cutting and P. Vishton, “Perceiving layout and knowing distance: the integration, relative potency and contextual use of different information about depth,” in Perception of Space and Motion, W. Epstein and S. Rogers, Eds., New York: Academic Press, 1995, pp. 69-118.

[9] A.E., Welchman, A. Deubelius, V. Conrad, H.H. Bülthoff, and Z. Kourtzi, “3D shape perception from combined depth cues in human visual cortex,” Nature Neuroscience, vol. 8, pp. 820-827, 2005.

[10] H. Sedgwick, “The geometry of spatial layout in pictorial representation,” in The Perception of Pictures, Volume 1: Alberti’s Window: The Projective Model of Pictorial Information, M. Hagen, Eds., London: Academic Press, 1980, pp. 34-90.

[11] J.M.Wallace and P. Mamassian, “The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces,” Vision Research, vol. 44, no. 19, pp. 2253-2267, 2004.

[12] F Metelli, “An algebraic development of the theory of perceptual transparency,” Ergonomics, vol. 13, no. 1, pp. 59-66, 1970.

[13] F. Metelli, “The perception of transparency,” Sci Am., vol. 230, no. 4, pp.:90-98, 1974.

[14] W. Gerbino, C.I. Stultiens, J.M. Troost, and C.M. de Weert, “Transparent layer constancy,” J. Experimental Psychology: Human Perception and Performance, vol. 16, no. 1, pp. 3-20, 1990.

[15] M. Zanforlin and L. Tommasi, “The perception of transparency with chromatic colours,” in Research in Perception, Logos, Padova, 1997, pp. 47-68.

[16] M. D’Zmura, P. Colantoni, K. Knoblauch, and B. Laget, “Color transparency,” Perception, vol. 26, pp. 471-492, 1997.

[17] F. Faul and V. Ekroll, “Psychophysical model of chromatic perceptual transparency based on substractive color mixture,” JOSA A, vol. 19, no. 6, pp. 1084-1095, 2002.

[18] B.G. Khang and Q. Zaidi, “Cues and strategies for color constancy: Perceptual scission, image junctions and transformational color matching,” Vision Research, vol. 42, no. 2, pp. 211-226, 2002.

[19] M. Singh and B.L. Anderson, “Toward a perceptual theory of transparency,” Psychological Review, vol. 109, no. 3, pp. 492-519, 2002.

[20] C. Ripamonti, S. Westland, and O. Da Pos, “Conditions for perceptual transparency,” J. Electronic Imaging, vol. 13, no. 1, pp. 29-35, 2004.

[21] B. Anderson, “The perceptual organization of depth, lightness, color, and opacity,” in The New Visual Neurosciences, J.S. Werner and L.M. Chalupa, Eds., Cambridge: MIT Press, 2014, pp. 653-664.

[22] J.M. Kraft and D.H. Brainard, “Mechanisms of color constancy under nearly natural viewing,” Proc. National Academy of Sciences, vol. 96, no. 1, pp. 307-312, 1999.

[23] T. Nagai, Y. Ono, Y. Tani, K. Koida, M. Kitazaki, and S. Nakauchi, “Image regions contributing to perceptual translucency: A psychophysical reverse-correlation study,” Iperception, vol. 4, no. 6, pp. 407-428, 2013.

[24] W. Kim, S. Ellis, M. Tyler, B. Hannaford, and L. Stark, “Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks,” IEEE Trans. on Systems, Man and Cybernetics, vol. 17, no. 1, pp. 61-72, 1987.

[25] B. Kappé, J. Korteling, and W. Van de Grind, “Time-to-contact estimation in a driving simulator,” in Proc. SID International Symposium, pp. 297-300, 1995.

[26] M.A. Kersten, A.J. Stewart, N. Troje, and R. Ellis, “Enhancing depth perception in translucent volumes,” IEEE Trans. Visualization and Computer Graphics, vol. 12, no. 5, pp. 1117–1124, 2006.

[27] C.R.C. Guibal, B. Dresp, “Interaction of color and geometric cues in depth perception: When does ‘‘red’’ mean ‘‘near’’?,” Psychological Research, vol. 69, pp. 30-40, 2004.