

Tamboli, Roopak R., Kara, Peter A., Cserkaszkzy, Aron, Barsi, Attila, Martini, Maria G., Jana, Soumya and (2018) "Canonical 3D object orientation for interactive light-field visualization", Applications of Digital Image Processing XLI, Proceedings of SPIE Volume 10752, ISSN 0277-786X (2018) <https://doi.org/10.1117/12.2320556>. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Canonical 3D object orientation for interactive light-field visualization

Roopak R. Tamboli^{a,c}, Peter A. Kara^b, Aron Cserkaszkzy^a,
Attila Barsi^a, Maria G. Martini^b, and Soumya Jana^c

^aHolografika, Budapest, Hungary

^bWMN Research Group, Kingston University, UK

^cIndian Institute of Technology Hyderabad, India

ABSTRACT

Light-field visualization allows the users to freely choose a preferred location for observation within the display’s valid field of view. As such 3D visualization technology offers continuous motion parallax, the users location determines the perceived orientation of the visualized content, if we consider static objects and scenes. In case of interactive light-field visualization, the arbitrary rotation of content enables efficient orientation changes without the need for actual user movement. However, the preference of content orientation is a subjective matter, yet it is possible to be objectively managed and assessed as well. In this paper, we present a series of subjective tests we carried out on a real light-field display that addresses static content orientation preference. The state-of-the-art objective methodologies were used to evaluate the experimental setup and the content. We used the subjective results in order to develop our own objective metric for canonical orientation selection.

Keywords: canonical view, subjective preferred orientation, light-field, 3D display

1. INTRODUCTION

Finding the best, preferred or canonical views of objects has been studied for several decades.¹ Knowledge of the canonical views or preferred orientations of 3D models helps several applications, such as automatic camera replacement, 3D scene generation,² scene navigation, scientific visualization, object recognition, mesh simplification, surgery planning, view-based 3D object recognition etc.³ Several studies⁴⁻¹⁰ have shown that orientation preference or view-canonicity is dictated by multiple factors such as: recognition, familiarity, functionality and aesthetic criteria. To seek preferred views, the researchers have adopted different methods, including ranking photographs taken from several viewpoints, imagining a viewpoint of an object based on its name followed by photographing the object to obtain same view, and orienting 3D models displayed on a 2D screen, using a controller.⁵ In this paper, we conducted a subjective study to find orientation preference when the 3D models are visualized on a light-field 3D display.

Light-field displays have opened new avenues in entertainment, scientific visualization, education etc.¹¹ Such displays provide glasses-free 3D experience via the virtually-continuous parallax effect. Visualization on such displays reveals significantly more information about the content than the 2D or stereoscopic displays. Not surprisingly, content generation for such displays poses a variety of challenges. Real-life content needs to be captured using multiple cameras maintaining very high angular resolution.¹² For small scenes — such as a single object — one can resort to turntable, setup enabling very high angular resolution while using a single camera.¹³ Knowledge of preferred views of objects being photographed can certainly help content acquisition. For instance,

Further author information: (Send correspondence to Roopak R. Tamboli)

Roopak R. Tamboli: E-mail: r.tamboli@holografika.com, ee13p0008@iith.ac.in

Peter A. Kara: E-mail: p.kara@kingston.ac.uk

Aron Cserkaszkzy: E-mail: a.cserkaszkzy@holografika.com

Attila Barsi: E-mail: a.barsi@holografika.com

Maria G. Martini: E-mail: m.martini@kingston.ac.uk

Soumya Jana: E-mail: jana@iith.ac.in

preferred views can be sampled more densely than the other views. Furthermore, composing a 3D scene using a computer graphics software can also be benefited by guiding the objects’ placement and orientation. To this end, we conduct a preliminary study to find preferred orientations of a few models visualized on a light-field display.

The rest of this paper is organized as follows: Section 2 reviews existing work on the preferred orientation of 3D objects. The methodology adopted in this paper is detailed in Section 3, together with the 3D models chosen for the experiments. Section 4 presents the results of our study. Section 5 discusses the results and their broader impact. The paper is concluded in Section 6.

2. RELATED WORK

The topic of what defines canonical views or preferred orientations has been studied for several decades. In an early investigation, Blanz *et al.* assessed the canonical views for objects by allowing participants to actively rotate realistically shaded 3D models.⁵ Objects were viewed on a 2D display and manipulated with a three-degree-of-freedom input device. In the first experiment, participants adjusted each object to the viewpoint from which they would take a photograph. In the second experiment, participants were asked to mentally image objects based on their name and then adjusted the object to the viewpoint from which they imagined it. Both the preferred views were found to be highly consistent for a given object.

Pertaining to view-based perception of 3D objects, Peters *et al.* provided several important insights, with focus on object recognition.¹⁴ They asserted that (i) Real-world objects possess canonical views, which are better suitable for recognizing the object than other non-distinguished views. (ii) The number of views which are sufficient to represent an object depends on the specific object. (iii) Certain view interpolation probably takes place, between familiar views when unfamiliar views have to be recognized.

Dutagaci *et al.* provided a benchmark that involved a methodology and a quantitative measure to evaluate the performance of view-selection algorithms.¹ They designed web-based subjective experiments to analyze humans’ view preferences and to construct a ground-truth. The performance of a view-selection algorithm was assessed with respect to the geodesic distance between the viewpoint determined by the algorithm and the viewpoint preferred by a human subject. Bonaventura *et al.* recently presented a survey of the computational measures that focus on the “goodness for recognition”.³ In their survey, they tested a common framework for the most basic measures introduced to select the best views for polygonal models.

In a recent work, Tamboli *et al.* observed that some primitive features of the visualized light-field content affected viewers’ experience, apart from the distortions added to the source stimuli.¹⁵ They observed that the viewing experience was affected not only by the spatial and angular quality of the content, but also by some seemingly primitive features, such as the area occupied by the respective object on the screen, part of the objects visible in a given 3D view etc. The differences in inter- and intra-3D view subjective ratings show that certain views of the objects were preferred over others, even when the available 3D views mainly depicted the frontal part.⁴ Their experiment can roughly be categorized as ranking of pre-selected 3D viewpoints.

To the best of knowledge of the authors, research on finding canonical views or preferred orientation was carried out exclusively on 2D displays. While such works provided algorithmic ways to predict preferred orientations, they have not been tested when the content is visualized on a glasses-free 3D light field display. To this end, in this exploratory paper, we present the a subjective study in which subjects oriented various 3D objects rendered on a light-field display according to their preference. In the next section, we discuss considerations towards designing an objective metric.

3. METHODOLOGY

In this paper, we chose 20 models from the ‘3D Model Data Set’ made available by Dutagaci *et al.*¹ The models are described in the Section 3.1. Details of subjective evaluation are described in Section 3.2.

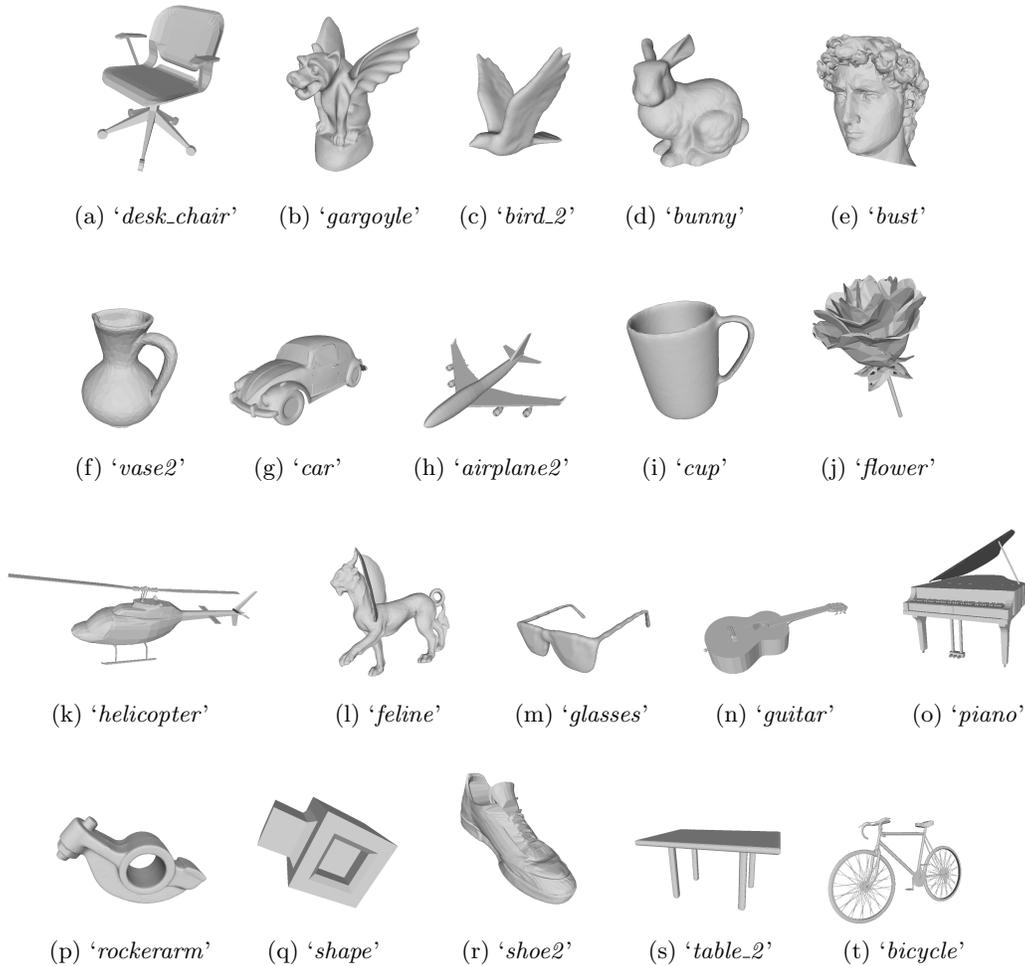


Figure 1: 3D models used in this paper.

3.1 3D Models

We chose 20 objects from the ‘3D Model Data Set’, made available by Dutagaci *et al.*¹ The ‘3D Model Data Set’ included models from ‘The Stanford 3D Scanning Repository*’ and the ‘SHape REtrieval Contest 2007: Watertight Models Track’.¹⁶

The models were selected from previously conducted studies primarily to be able to follow the existing framework. Since this study was first-of-its-kind to the best of our knowledge, the selected models mainly contained simple well-known objects encountered in real-life. In the following, we categorize the selected models, referred by their names in the ‘3D Model Data Set’.

1. Well-known objects encountered in daily life: ‘airplane2’, ‘bicycle’, ‘bird_2’, ‘bunny’, ‘car’, ‘desk_chair’, ‘cup’, ‘flower’, ‘glasses’, ‘guitar’, ‘helicopter’, ‘piano’, ‘shoe2’, ‘table_2’, ‘utah_teapot’, ‘vase2’, ‘wine_glass’.
2. Artistic objects: ‘gargoye’, ‘feline’, ‘bust’.
3. Previously unknown or abstract models: ‘rockerarm’, ‘shape’.

*<http://graphics.stanford.edu/data/3Dscanrep/>



Figure 2: Subjective evaluation.

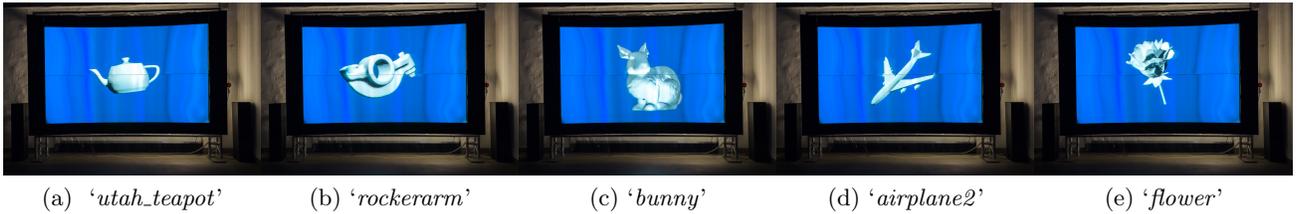


Figure 3: Photographs of 3D models visualized on the C80 cinema system.

3.2 Subjective Evaluation

The subjective tests were carried out in an isolated laboratory environment, as shown in Figure 2. The stimuli were presented on the HoloVizio C80 light-field cinema,¹¹ which is a front-projection system with a 3-meter-wide screen. The FOV of the system was calibrated for 45 degrees. The test participants viewed the screen of the light-field display from a distance of $2.5H$, which corresponded to 4.6 meters. The objects were rendered using a software called ‘ContentViewer’. The order as well as initial orientation of 3D models were randomized for each subject. The software allows translations, rotations and scaling of the displayed model. One can apply the said transformations to either the object or the camera. In this work, subjects were allowed to rotate the objects, using thumbstick on a Logitech F310 gamepad[†]. The test subjects were given sufficient time to try out various orientations of the displayed object. Subjects were also allowed to move sideways in front of the display, within the FOV. To support observer movement, the USB cable of the gamepad — connected to the computer controlling the display — was extended. Once a preferred view was found, subjects notified the display operator to save their preference. In this manner, preferred orientations for 20 objects were collected per test participant. The subjective test required approximately 20 minutes. Since the 3D object was at the center of the screen and no scaling was allowed, a preferred orientation was characterized by 3 rotations. In this study, 10 subjects provided their preferred orientations.

4. RESULTS

We now present the results of our subjective evaluation. Figure 4 shows the representative orientations of some of the models. The orientation preferences supplied by subjects for all these models can be derived by rotating the models about the Y-axis. Due to the horizontal parallax, all those views were always available to all the subjects upon moving in front of the display and therefore only one of them is shown in Figure 4. For most of the common models, the preferred orientations are front views with some elevation angle. In particular, the models ‘bunny’, ‘bust’, ‘chair’, ‘cup’, ‘bicycle’, ‘utah.teapot’, ‘vase2’, ‘glasses’, ‘helicopter’ and ‘table’ seem to have canonical views as all the subjects preferred nearly the same orientation. Among the remaining models, preferred orientations for the well-known models, namely ‘bird_2’, ‘car’, and ‘shoe’, as well as unfamiliar models, namely, ‘rockerarm’ and ‘shape’ seem to reveal maximum information about those objects.

[†]<https://www.logitechg.com/en-us/product/f310-gamepad>

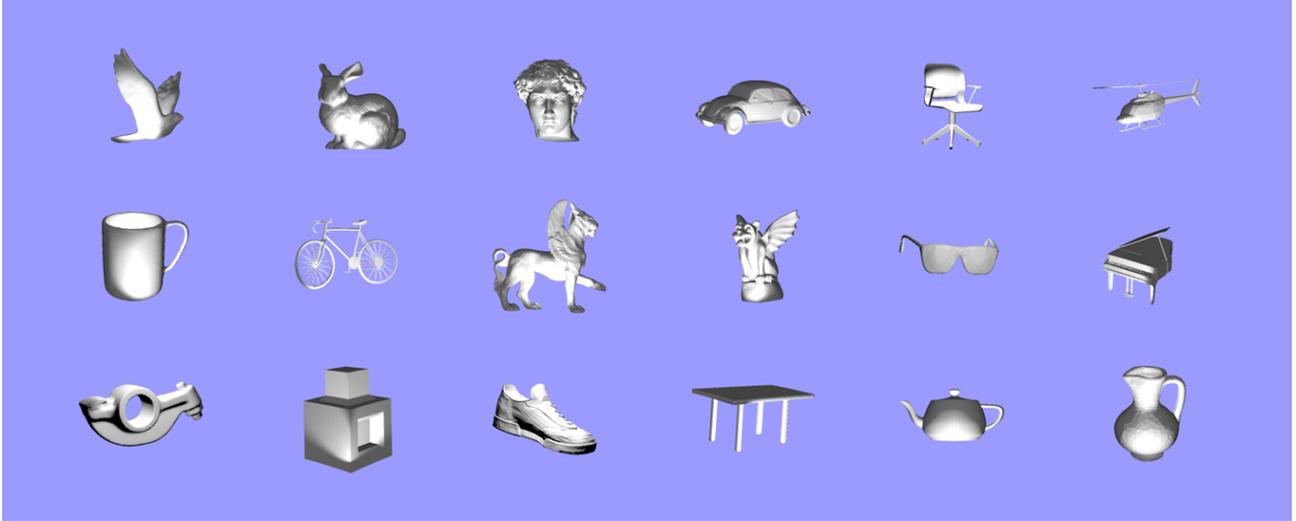


Figure 4: 3D models for which all preferred orientations could be seen via horizontal parallax.

On the other hand, orientation preferences for three models, namely ‘*flower*’, ‘*airplane2*’ and ‘*guitar*’ were varied across subjects. Some of the orientations are shown in Figure 5. These three models have been separated from the others because the preferred orientations included rotations about all the three axes. Since the light-field display provided a realistic experience, orientations which are unlikely to be chosen on a 2D displays, such as a plane with its nose pointing out from the display, were also preferred.

4.1 Objective Metric Considerations

While the results of the subjective evaluation study corroborate in general with the observations reported by previous works,^{1,5} the objective evaluation in case of light-field visualization is more involved. The subjects evaluated an orientation from multiple viewing locations in front of the display which enabled access to a very large set of 2D views. Since the available objective metrics evaluate distinct 2D views, such metrics may not be directly applicable to the 3D content. In a recent attempt, Tamboli *et al.* were able to propose image-based metrics on light-field content because the 3D views in their work were composed of a set of 2D views and subjects did not interact with the content.¹⁵

Datugaci *et al.* evaluated seven view selection algorithms.¹ Those were (i) View area (same as Tamboli *et al.*), (ii) Ratio of visible area, (iii) Surface area entropy, (iv) Silhouette length, (v) Silhouette entropy, (vi) Curvature entropy, and (vii) Mesh saliency. ‘Surface area entropy’ and ‘Silhouette length’ resulted in the least average view selection error, when compared to their subjective scores. These algorithms, as said earlier, were evaluated on per-view basis which computed the goodness of a view utilizing projections of meshes. One possible way of using these measures for light-field visualization could be the pooling of several per-view scores. However, the number of views to be considered — i.e., discretization on the FOV — is a concern. More discretization in complex areas of models and less discretization in simpler areas could be a solution. The FOV of the display could simplify the task to certain extents. One can render light-field slices supplied to the projectors of the display, which are usually very different than usual 2D images, and apply image-based metrics on those.¹⁷ However, such images are still different than what an observer would see in front of the screen of the display.

5. DISCUSSION

The 20 objects used in this study included objects encountered in daily life, artistic objects, as well as unfamiliar objects. In general, we found that the preferred orientations for well-known objects were very close to each other and were dictated by their functionality and real-world appearances. The subjects oriented the artistic objects in an aesthetically appealing manner. On the other hand, the two unknown objects were adjusted such that

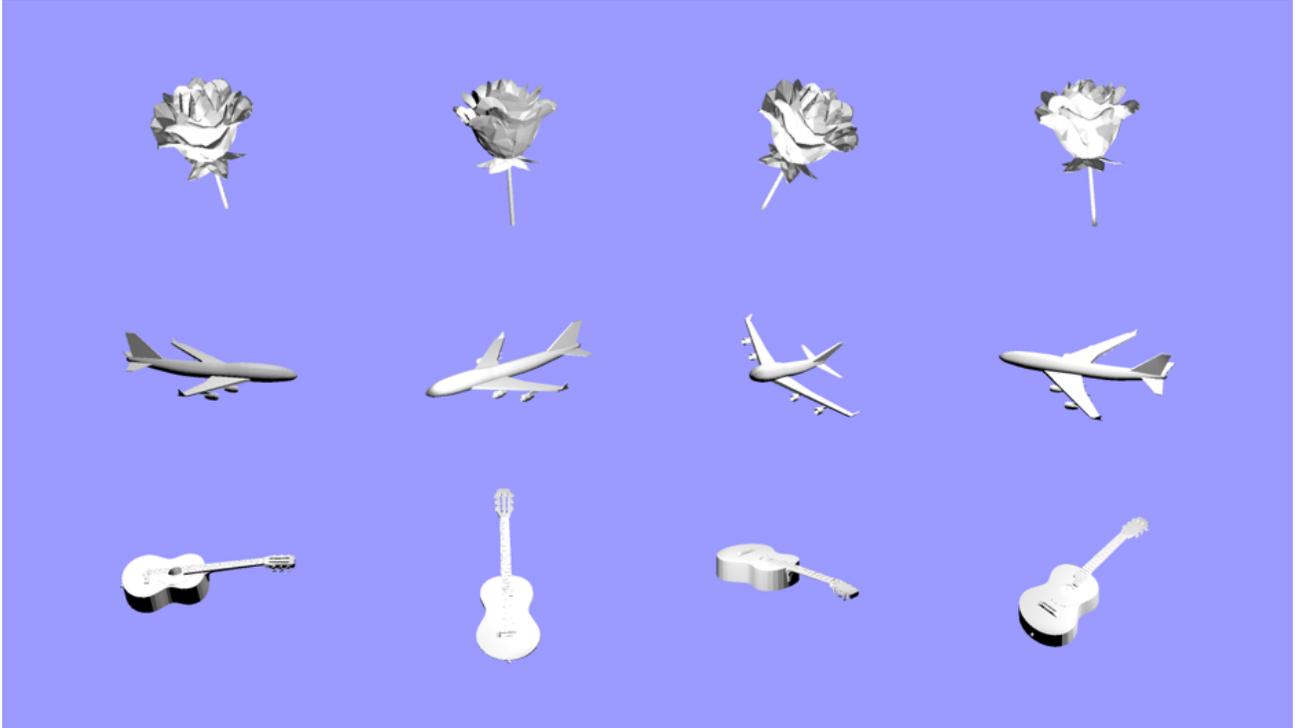


Figure 5: 3D models for which all preferred orientations could not be seen via horizontal parallax.

maximum information about them was revealed. Furthermore, it was possible to receive an unstable view; a view which changes drastically upon small changes, from certain viewing locations in front of the display. Although this issue was not envisaged in this work, the test subjects preferred the orientations which did not include such unstable views from any possible viewing location. For instance, no subject preferred an uncommon orientation such as bottom of a chair or an unstable orientation such as the nose of the fuselage of the plane.

Similarly to study conducted by Blanz *et al.*,⁵ our subjective evaluation methodology provides: (i) virtually infinite viewpoints, (ii) direct measurement of viewing angles by the virtue of the rendering software used, (iii) active exploration of viewpoints. While these features are easily available nowadays, the most distinctive feature was the visualization of a light-field display.

Since our study was conducted on a glasses-free light-field display, only one preferred orientation was sought. For any given orientation, most of the next preferred views (which appeared by rotating object about Y-axis) were already available to the subjects due to the virtually continuous horizontal parallax offered by the display. In other words, it can be said that a preferred orientation chosen by a subjects provided a cluster of several candidate 2D views, which would have been preferred by the same subject on a 2D display. The mode of interaction with the objects was also very intuitive which required little or no training.

However, compared to various other subjective evaluations conducted on light-field displays, seeking preferred views is highly challenging. For instance, Kara *et al.* studied the impact of a few combinations of spatial and angular resolutions of the light-field content, and thereby investigated the interplay between them, on the perceived quality of light-field visualization.^{12,18} Tamboli *et al.* devised a full-reference spatio-angular objective quality metric to study effects of a few spatial distortions on the perceived quality.¹⁹ The stimuli used in these and other similar studies were carefully controlled in order to examine a few factors at a time which affect viewing experience on light-field displays. On the other hand, seeking preferred orientation allows very little control over the outcome of such studies because of various criteria that subjects use to evaluate a view, unless explicitly told.

6. CONCLUSION

In this work, we presented a pilot study on seeking preferred orientation of 3D objects on a glasses-free light-field display. This preliminary exploration was carried out with 10 subjects who supplied their preferred orientation for 20 3D models. For majority of the preferred viewpoints of the objects were very close. We plan to include a wider and diverse set of 3D models in our future studies. We also plan to provide view selection criteria to subjects and also include a post-study questionnaire.

ACKNOWLEDGMENTS

The work in this paper was funded from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreements No 676401, European Training Network on Full Parallax Imaging and No 643072, Network QoE-Net.

REFERENCES

- [1] Dutagaci, H., Cheung, C. P., and Godil, A., “A benchmark for best view selection of 3d objects,” in [*Proceedings of the ACM Workshop on 3D Object Retrieval*], *3DOR ’10*, 45–50, ACM, New York, NY, USA (2010).
- [2] Wong, L. M., Dumont, C., and Abidi, M. A., “Next best view system in a 3d object modeling task,” in [*Computational Intelligence in Robotics and Automation, 1999. CIRA ’99. Proceedings. 1999 IEEE International Symposium on*], 306–311 (1999).
- [3] Bonaventura, X., Feixas, M., Sbert, M., Chuang, L., and Wallraven, C., “A survey of viewpoint selection methods for polygonal models,” *Entropy* **20**(5), 370 (2018).
- [4] Tarr, M. J. and Kriegman, D. J., “What defines a view?,” *Vision Research* **41**(15), 1981 – 2004 (2001).
- [5] Blanz, V., Tarr, M. J., and Blthoff, H. H., “What object attributes determine canonical views?,” *Perception* **28**(5), 575–599 (1999).
- [6] Massios, N. A. and Fisher, R. B., “A best next view selection algorithm incorporating a quality criterion,” in [*Proceedings of the Ninth British Machine Vision Conference*], 780–789, Informatics Institute (IVI) Institute for Logic, Language and Computation (ILLC), Southampton: British Machine Vision Association (1998).
- [7] Niimi, R. and Yokosawa, K., “Three-quarter views are subjectively good because object orientation is uncertain,” *Psychonomic Bulletin & Review* **16**(2), 289–294.
- [8] Secord, A., Lu, J., Finkelstein, A., Singh, M., and Nealen, A., “Perceptual models of viewpoint preference,” *ACM Trans. Graph.* **30**, 109:1–109:12 (Oct. 2011).
- [9] Carrasco, M., “Visual attention: The past 25 years,” *Vision Research* **51**(13), 1484 – 1525 (2011). Vision Research 50th Anniversary Issue: Part 2.
- [10] Mortara, M. and Spagnuolo, M., “Semantics-driven best view of 3d shapes,” *Computers & Graphics* **33**(3), 280 – 290 (2009). IEEE International Conference on Shape Modelling and Applications 2009.
- [11] Balogh, T., Nagy, Z., Kovács, P. T., and Adhikarla, V. K., “Natural 3D content on glasses-free light-field 3D cinema,” in [*IS&T/SPIE Electronic Imaging*], 86480F–86480F, International Society for Optics and Photonics (2013).
- [12] Kara, P. A., Cserkaszkzy, A., Darukumalli, S., Barsi, A., and Martini, M. G., “On the edge of the seat: Reduced angular resolution of a light field cinema with fixed observer positions,” in [*Ninth International Conference on Quality of Multimedia Experience (QoMEX)*], (2017).
- [13] Tamboli, R. R., Reddy, M. S., Kara, P. A., Martini, M. G., Channappayya, S. S., and Jana, S., “A High-angular-resolution Turntable Data-set for Experiments on Light Field Visualization Quality,” in [*Quality of Multimedia Experience (QoMEX), 2018 Tenth International Conference on*], IEEE (2018).
- [14] Peters, G., [*A View Based Approach to Three Dimensional Object Perception*], Shaker (2002).
- [15] Tamboli, R. R., Kara, P. A., Appina, B., Martini, M. G., Channappayya, S. S., and Jana, S., “Effect of Primitive Features of Content on Perceived Quality of Light Field Visualization,” in [*Quality of Multimedia Experience (QoMEX), 2018 Tenth International Conference on*], IEEE (2018).
- [16] Giorgi, D., Biasotti, S., and Paraboschi, L., “Shape retrieval contest 2007: Watertight models track.”

- [17] Holliman, N. S., Dodgson, N. A., Favalora, G. E., and Pockett, L., “Three-Dimensional Displays: A Review and Applications Analysis,” *IEEE Transactions on Broadcasting* **57**, 362–371 (Jun. 2011).
- [18] Kara, P. A., Cserkaszkzy, A., Barsi, A., Papp, T., Martini, M. G., and Bokor, L., “The interdependence of spatial and angular resolution in the quality of experience of light field visualization,” in [*2017 International Conference on 3D Immersion (IC3D)*], (Dec 2017).
- [19] Tamboli, R. R., Appina, B., Channappayya, S., and Jana, S., “Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display,” *Signal Processing: Image Communication* **47**, 42–55 (2016).