Eurographics Workshop on 3D Object Retrieval (2026)
S. Biasotti, E. Moscoso Thompson, and M. Spagnuolo (Editors)

Harvesting natural points-of-view preferences for arbitrary objects

Gianluca Roveda[1], Georgios Papaioannou[2], Andreas A. Vasilakis[2], Marco Tarini[1]

[1] University of Milano, Italy
[2] Athens University of Economics and Business, Greece

Abstract

Automatically identification of an optimal, representative or relevant viewpoint for a given 3D object is an important task with several applications in digital marketing, visualization, 3D data management and shape retrieval. An objective function to optimize the viewpoint is difficult to define in purely geometric terms, given the semantic and aesthetic bias a user is likely to introduce when presented with a target object. Therefore, supervised data-driven approaches are a natural candidate to address the problem. This implies the availability of large datasets describing natural point-of-view preferences for a number of object categories. In this work, we present a framework designed to harvest this kind a dataset. The system involves tasking end-users with answering general qualitative questions about a displayed 3D object, while silently collecting point-of-view information. In other terms, we capture viewpoint preferences as a latent objective, allowing for an unbiased observation of user behavior in performing various tasks that involve object visualization. We publicly release both our dataset of collected viewpoint statistics and the tool to harvest them, which is easily configurable by experimenters to collect more data in the future or integrate the approach into online 3D viewers.

Keywords: viewpoint determination, object inspection, visual analytics

1. Introduction

The ability to choose a viewpoint for an object in such a way that effectively communicates its functional and aesthetic aspects to the target observer, without the use of additional explanatory elements (such as annotations or captions), remains an under-explored area, despite its potential impact in various fields, from visual communication, to human-machine interaction [Nor13], with direct applications ranging from e-commerce, thumbnail creation, shaper retrieval, or even 3D scene navigation [BBT*13].

The viewpoint from which an object is observed has a decisive influence on visual perception: it may increase the attractiveness of the object and make it more easily recognizable, or conversely, it may cause ambiguity, confusion and loss of focus to important details. This becomes especially crucial for 2D renderings of three-dimensional objects and environments, such as a photorealistic rendering for catalog display, the production of promotional content and 2D or 3D presentation of objects in online shops. The current strategy is to use images from different standard angles, or to choose the angle according to personal taste.

In some object categories, the optimal pose can be intuitive or trivial to suggest, like when the objects are of simple, symmetrical structure or when is a predefined artistic intent in their presentation. In most cases, determining a well-placed projection of an object is more difficult. Many cases arise when subjective goals are involved, such as looking for a particular detail of interest, or when factors external to the object itself are involved, including relative positioning within a virtual environment or interaction with lighting.

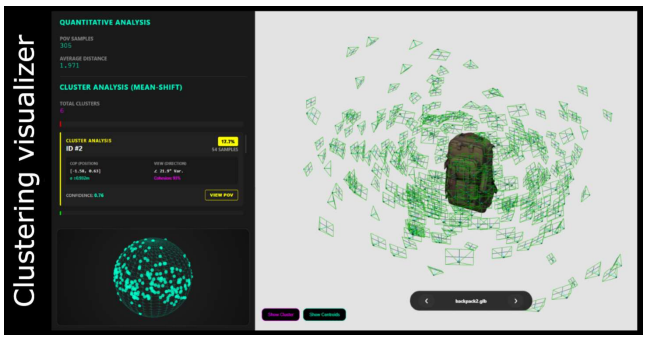

Figure 1: A snapshot of our dataset analyzer. The sidebar aggregates session statistics, such as total sample count and average interaction distance. The Mean-Shift clustering is shown through individual data cards a Polar Heatmap on a wireframe sphere shows the 3D viewpoint distribution (lower right).

Previous studies have addressed the problem by taking a geometric approach, based on shape analysis, in which the determination of the optimal viewpoint is made by metrics derived directly from the structure of the 3D model. The proposed methods calcu-

late visual entropy based on the distribution of visible surfaces, or assess the local saliency of vertices by weighting their exposure in a given view [LVJ05]. However, other studies have highlighted how visual perception is deeply influenced by cognitive and emotional mechanisms [BTB99]. In fact, research in neuroscience and psychology shows that different shapes can cause different affective reactions [BN07].

It is therefore natural to ask: what criteria should be used to select the viewpoint? Should a purely geometric approach be favored, maximizing visibility of surface area or salient features? Or is it more appropriate to highlight the functional features of the object, those that tell its purpose or function? How do aesthetic factors blend in with the functional aspects of the object? Which categories of objects are more prone to subjective preference of a viewpoint? Can we identify clusters of preferential viewing directions? These questions open the way for a broader reflection on the relationship between object, observer, and visual representation, suggesting the need for more user-centered approaches.

The motivation behind this work lies in the desire to approach the problem through direct observation of how users spontaneously interact with 3D objects in digital environments. The goal is to develop a data-driven approach to continuously collect and analyze data about the behavior of observers to deduce the most meaningful viewpoint(s). Furthermore, users often approach the same object with different goals. For example, inspecting its overall structure, examining specific parts, or understanding functional aspects. These goals can lead them to follow different exploration paths in the viewing space. Consequently, the viewpoints they select or spend more time observing are influenced not only by the object's semantics but also by the users' underlying intent, which, in contrast to previous works, becomes a key factor in our investigation.

This paper contributes to the field of visual analytics and object retrieval as follows:

- we introduce the idea to base viewpoint selection on actual observations of user interactions with 3D objects;
- we develop a tool to harvest user's preferences of viewpoints: importantly, viewpoints are collected indirectly, without explicitly asking users for a preferred viewing angle, but observing their choices while they perform a 3D task;
- we provide a first harvested dataset, offer a preliminary visual and quantitative analysis, and offer a framework to repeat this analysis.

2. Related Work

The problem of determining the optimal view for a given 3D model is often termed the "best view" problem. We refer to a recent survey [BFS*18]. In the following, we outline more relevant contributions.

2.1. Geometric approaches

Several approaches have been proposed to define the optimal viewpoint focusing on geometric criteria.

Earlier methods determined the validity of a viewpoint a function of the visibility of the faces of the model. Viewpoint entropy [VFSH03], for example, is a measure of visual information calculated from the distribution of projected areas of faces visible from a given viewpoint. The assumption is that a view is more informative if it has a well-balanced, heterogeneous distribution of visible surfaces. This is extended by incorporating mutual information between different views [FSG09], to guide the viewpoint selection in multi-view contexts.

More recent approaches assess the saliency of each vertex of the 3D model, a measure of how much a point on the surface stands out compared to its neighbors; potential viewpoints are then scored according to the total saliency of the visible vertices [NCL15], optionally favoring frontal views of salient points [LST12].

Our data-driven approach requires no assumption about the relevance of a-priori geometrically justified measures such as saliency or view entropy.

2.2. Beyond geometry

Several studies suggest that shape perception is not explained solely by geometric terms. Different shapes, such as curved or angular forms, evoke distinct emotional responses at a neural level, indicating a deep link between geometry and emotions [BN07, BPGG16]. Similarly, geometric illusions reveal that perception is influenced by internal psychological mechanisms, beyond mere spatial measurement [SSC16]. According to Gestalt psychology, perceptual organization occurs on the basis of mental and subjective principles that actively structure visual experience [WEK*12].

Neural approaches use a limited number of pre-defined viewpoints and map the problem as a selection process, from which preferential views are inferred [HSO*22]. In contrast, we use free viewpoint exploration and continuous recording.

Models based on a linear combination of visual criteria have been proposed [SLF*11]: visible area, contours, and salient edges, which are assigned differential weights learned from experimental data. This approach not only considers geometric information, but aims to reflect observers' preferences, thus approaching shape evaluation based on subjective visual experience derived from asking subjects to choose the best views of 3D models.

Semantics-driven approach [MS09, Lag10] extracting semantic features through geometric analysis or structural decomposition of the 3D mesh, and assume that good viewpoints result in good representation and balance of the semantically distinct areas in the resulting image. Again, our purely data-driven approach is complementary and completely foregoes that assumption.

A seminal contribution is the benchmark by [DCG10], which formalizes the discrepancy between purely geometric metrics and human judgment through a large-scale user study. Although their work establishes a ground truth for human-centric selection, our approach leverages this principle of collective preference to build a more streamlined and accessible data-driven process.

Despite its limited scope, a recent study [JMF*25] analyzes how users in VR environments select preferred viewpoints for 3D graphs. Preferences emerge from aesthetic measures such as visual stress, node overlap, and principal axis visibility; these are not ge-

ometry based, but subjective evaluations driven by layout and visual clarity.

These approaches are close to the direction of our work. Our proposal differs in the adopted method: instead of using complex technologies such as virtual reality or conducting elaborate studies that integrate geometric and psychological factors, we rely on a simpler and more accessible approach, guided solely by the collected data the identification of the best viewpoints. More importantly, in our questionnaire design, we keep the view-selection objective undisclosed, aiming at a more natural observation of the model inspection process.

3. Methodology

The proposed system consists of a *harvesting tool* to capture preferred viewpoints (Sec. 3.1) as well as a *visualization tool* to preview and analyze captured viewpoints (Sec. 4.1).

Using the data harvesting tool, *test subjects*, the participants of the experiment, interactively inspect a 3D model, while the system silently captures their viewpoint preferences as they perform seemingly unrelated quizzes on the 3D model. Its interface is designed for clear and intuitive interaction, facilitating the natural exploration of the 3D models, even for completely inexperienced users.

Before the experiment, a *data analyst* prepares the set of questions, by defining and customizing each question in the questionnaire through a set of parameters. Our tool includes an easy to use framework for this task. Unlike for test subjects, we assume a basis of computer science literacy from the data analysts preparing the questionnaire.

The entire system is implemented as a web-application that can be hosted in a web site, to facilitate ease of access and facilitate the administration of the questionnaire.

3.1. Data Capture Tool

In our system, the capture of viewpoint data is performed indirectly, by displaying 3D models of objects and challenging the participants to perform various object-related tasks that require the inspection of the objects with different intents The tasks are provided in the form of a questionnaire can include questions such as *"Which of these T-shirts do you like best?"*. Users interact with the 3D models in order to provide an answer, guided by their personal aesthetic preferences and visual perception of the object's features or functionalities. The system then records the used viewpoints, as a proxy for their visual interest. The input-output flow is designed to be non-intrusive, distinguishing between active navigation and intentional observation. Continuous camera movements allow for spatial exploration, and the system captures a viewpoint only when a stable pause is detected. These pauses are interpreted as indicators of visual interest, suggesting that the participant has found a viewpoint worthy of closer inspection.

In each session, we also collect some basic information about the participants. These data are provided voluntarily, anonymously, and only after the user has given their explicit consent. We include this information in the collected database, for any subsequent analysis.

3.1.1. Questionnaire Structure and Task Design

The questions are structured to promote active exploration and interaction with the 3D models. To observe how stylistic, structural, and geometric variations affect viewpoint selection, the question can include intra-class repetitions (presenting multiple distinct variations of items such as shoes, chairs, or cups).

Our framework supports three types of questions:

- **Multiple meshes:** The user is asked to compare and choose one out of several concurrently or sequentially shown 3D objects that best satisfies a specific query (see Fig. 2, left);
- **Single mesh:** A single 3D model is displayed, and the user must provide an answer linked to the object's attributes, perceived functionality, etc (see Fig. 2, top-right).
- **Photo mode:** The user freely explores a single mesh and manually triggers the capture to record the viewpoint they perceive as the most representative or aesthetically pleasing for taking a photograph (see Fig. 2, bottom-right). In this case, the automatic viewpoint capture is disabled.

To systematically investigate how user intent shapes observation behavior, the tasks can be formally coupled with the objects based on three semantic and functional dimensions:

1. **Functional Affordance & Semantics:** Applied primarily to single meshes, these queries stimulate the user to decode the object's purpose or utility (e.g., *"What does this machine train?"*, *"What is this knife suitable for?"*, or *"Where would you place this piece of furniture?"*).
2. **Context-Driven Selection:** These tasks force users to inspect the geometric and semantic features of multiple intra-class models to evaluate their situational appropriateness (e.g., *"Which shoe is more suitable for running?"*, *"Which backpack would you take to the mountains?"*, or *"Which car looks faster to you?"*).
3. **Subjective Hedonic Preference:** Designed to explore aesthetic judgment and personal taste, these questions capture potential hidden consensus in viewpoint preferences driven by emotional or commercial value (e.g., *"Which cup would you prefer as a gift?"*, *"Which armchair would you sit on to read a book?"*, or *"Which plate do you prefer?"*).

By hiding the viewpoint enquiries behind intuitive real-world decisions (utility, context, and taste), the system captures how humans naturally look at objects when actively decoding their functional and semantic meaning, ensuring the recorded data is free from the cognitive biases of explicit viewpoint elicitation tasks.

3.1.2. Triggering Viewpoint Sampling

To ensure geometric consistency, object coordinates are by default centered and normalized based on the bounding sphere of the object. The normalization matrices (representing the translation and scaling) are stored alongside each mesh, allowing the viewpoint to be mapped back to the original world space. All recorded camera parameters specifically the position $\mathbf{c}$, the viewing direction $\mathbf{f}$, and the up vector $\mathbf{u}$ are expressed relative to this normalized object space. Camera positions are not restricted to the normalized bounding sphere.

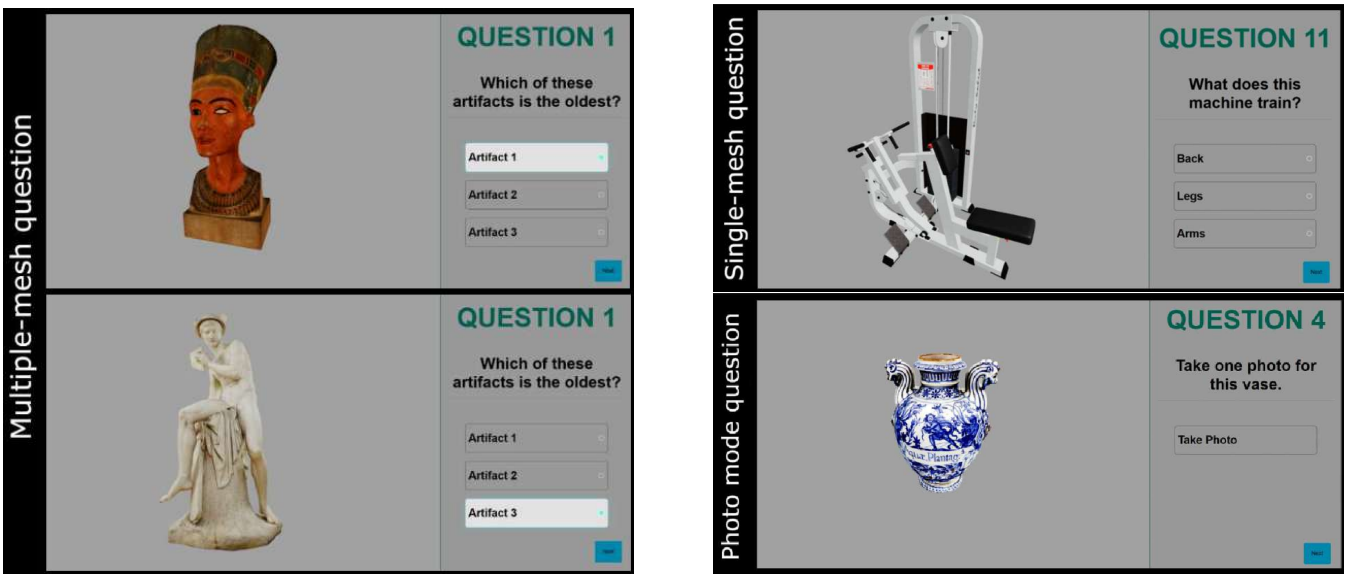

Figure 2: An example for each of the three supported types of questions.

Additionally, a recording is made to capture the final viewing position before switching to a new mesh. This is because we believe that, in most cases, users tend to leave the object in a stable position after gathering the necessary information.

The conditions that trigger the capture of a viewpoint are controlled by two threshold values: the stability time threshold t_t and the angular threshold t_α. Whenever the camera view-direction or up-vector stays within t_α angle for more than t_t seconds, a new viewpoint is captured; the same viewpoint is not captured again until the cumulated change exceeds t_α again. We empirically identified $t_\alpha = 7^\circ$ and $t_t = 0.7s$ to work well in practice.

3.1.3. Ethics and Privacy

To ensure that the collected data is free from cognitive bias, we do not initially reveal the actual purpose of the study. This is a form of controlled deception that conforms to major guidelines on ethical research with human subjects such as the Belmont Report [Nat79] and the ethical code of the American Psychological Association [Ame17].

In the debrief at the end of the experience, we explain the true purpose of the study and the type of data collected. At this stage, the test subject has full discretion to withdraw from the study and request deletion of their data, even if it was collected anonymously. This way, we maintain scientific integrity while protecting participants' rights and freedom of choice.

In order to comply with data protection regulations, the system is designed to operate completely anonymously. Each user is identified by a UUID (Universally Unique Identifier) generated at the beginning of the session. No personally identifiable information such as name or email is required or stored. The only demographic data collected is the age range (1–15, 16–25, 26–40, 41–65, 65+), potentially usable for future statistical analysis yet insufficient to identify individual identities.

3.2. Implementation

Our framework is designed to make it easy to define of new questioners over 3D objects. The succession of questions is represented as an array, each entry being a tuple with a title, the text of the question, a set of contextual parameters, the initial viewpoint, the type of camera controls (Sec. 3.2.1), and lighting parameters.

The framework adopts a client-server scheme, consisting of three main components:

Front end: an interactive web application that handles 3D model display, application submission, and user motion capture. 3D rendering and camera control are managed through Three.js;
Application server: a RESTful server developed in Python that receives and stores data in the database;
Database: a PostgreSQL storage system where the experiment data are recorded.

3.2.1. Camera Control

Although more sophisticated alternatives exists [MCS16], the system supports two simple mouse-controlled interaction models: Orbit and Trackball controls. Both modes allow users to control rotation, zooming, and panning. The Orbit mode constrains the camera's up-vector to the world Y-axis, preventing the object from appearing upside down and providing a more stable and intuitive experience for most users. In contrast, the Trackball mode offers unconstrained rotation around all axes, which is essential for inspecting objects that lack a clear "up" orientation or have poorly initialized coordinate systems. Experimenters can select the mode that best fits the specific 3D model or task to minimize unwanted movements and user frustration. The system also includes a flag to indicate whether the camera parameters should be restored to the last user-selected viewpoint after replacing a mesh, ensuring visual consistency or a restart as required by the context of the question.

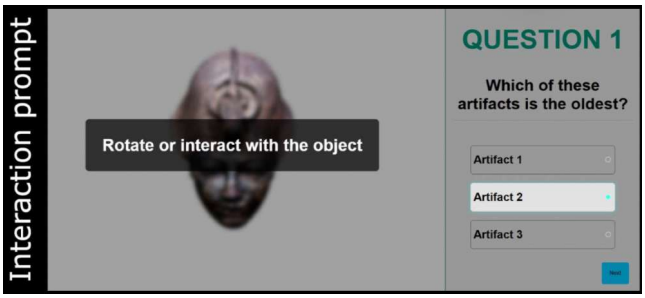

Figure 3: Interaction prompt that appears each time a mesh is loaded for the first time.

3.2.2. Motivating User Interaction

An issue identified during early testing was that when a model was loaded in a frontal position, many users interpreted it as a static image, drastically reducing their desire to interact with it. An initial attempt to address this behavior involved randomizing the initial camera angle when the mesh was loaded. However, this solution introduced undesirable effects, such as a lack of consistency between questions. Therefore, this strategy was replaced by a more controlled approach: we added a blur effect to the scene upon display of the content for the first time, accompanied by an informational message asking the users to interact with the model to continue. This solution had a strong positive impact, especially on less experienced users (see Fig. 3).

To further encourage interaction and ensure more fair coverage of the provided 3D models, we introduced a controlled progress system within the questionnaire: In order to answer a question, the participant must first interact with all the meshes associated with it. This constraint ensures genuine exposure to the models and helps prevent superficial or insufficiently explored responses.

4. Experiments

We tested our own system by conducting a first viewpoint harvesting session, involving approximately 50 participants. We implemented a questionaire consisting of 17 distinct questions distributed across a heterogeneous dataset of 45 3D meshes, harvesting a total of 10,006 user-generated viewpoints.

The 3D models were obtained from the Sketchfab platform [Ske26]. To ensure an engaging experience and minimize user fatigue, they represents a diverse range, including both realistic objects obtained through photogrammetry and artistic creations.

The results indicate that the interface is intuitive, with a high degree of task completion. After experiment discussion with testees revealed that the questioning system acted as an incentive for deeper object exploration: the participants' desire to provide 'correct' or thoughtful answers led to a meticulous examination of the 3D models. This 'gamified' pressure reduced casual or hurried interactions, making the captured viewpoints reflect a genuine, deliberate analysis of the functional or aesthetic features. We performed a preliminary data analysis on the captured data, as follows.

4.1. Methods of analysis of captured data

We compute aggregated statistics for each object: the total *Point of View* (POV) count and the *Average Distance* $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} \|P_i\|$, reflecting the users' preferred zoom level. To visualize distribution of view directions, a *Polar Heatmap* serves as a 3D spatial histogram: each POV is projected onto the unit sphere, highlighting (see Fig. 1, lower left).

Then, we also implemented a Mean-Shift clustering (which, unlike k-means, requires no *a priori* specification of the number of clusters k). See Fig 4 for an example.First, we remove zoom bias by projecting each view position into a unit sphere centered at the object's origin. Then, we fix a bandwidth parameter of $h = 0.5$, corresponding to a $\approx 30^\circ$ aggregation window. To avoid oversmoothing semantic viewpoints (and also to improve efficiency), we constrain the optimization to 5 iterations, which serves as an early-stopping regularizer. After convergence, we merge similar seeds with a merge threshold $\approx 11.5^\circ$. We remove outliers from erratic camera movements discarding any clusters including fewer than 4% of the viewpoints.

For each resulting cluster of n_c view directions $\mathbf{d}_i$, we extract the angular standard deviation σ_θ with Von Mises-Fisher:

$$\sigma_\theta = \sqrt{-2\ln\left(\left\|\frac{1}{n_c}\sum_{i=1}^{n_c}\mathbf{d}_i\right\|\right)}. \quad (1)$$

To foster further experiments and reproducibility, each of the above functionalities have been implemented in a reusable system that we provide as a part of our framework, complete with a user interface to compute and report the above statistics (see Fig. 1).

4.2. Results of the Statistical Analysis

The metrics extracted, the number of clusters k (see Fig. 5), the Main Cluster Percentage (*MCP*), and the Convergence Index (*CI*), revealed distinct categories of user observational behavior, apparently influenced by geometric shape, functional relevance, and textural information. By cross-referencing the statistical metrics (Table 1) with the visual evidence in Figures 6-8, three cases of user's behavior emerged:

Viewpoint consensus. A portion of the dataset exhibits a "unimodal" distribution, where users reach a consensus on the object's representative view. This *semantic frontalism* is particularly evident when the model has a clear functional orientation. For example, in the `wardrobe` example (Fig. 6-mid), participants instinctively adopted an interaction-oriented perspective, reflecting the typical angle of observation in a physical environment. For objects where a specific function is not the primary interest, such as the `statue`, users gravitate toward sides with the most salient features (Fig. 6-top), resulting in sharp density peaks and *CI* values exceeding 89.0. Furthermore, textural information appears to compete with, and sometimes override, geometric saliency: in the `mug` data (Fig. 6-bottom), the majority of viewpoints were driven by recognizable patterns on the textured surface rather than the handle's geometry.

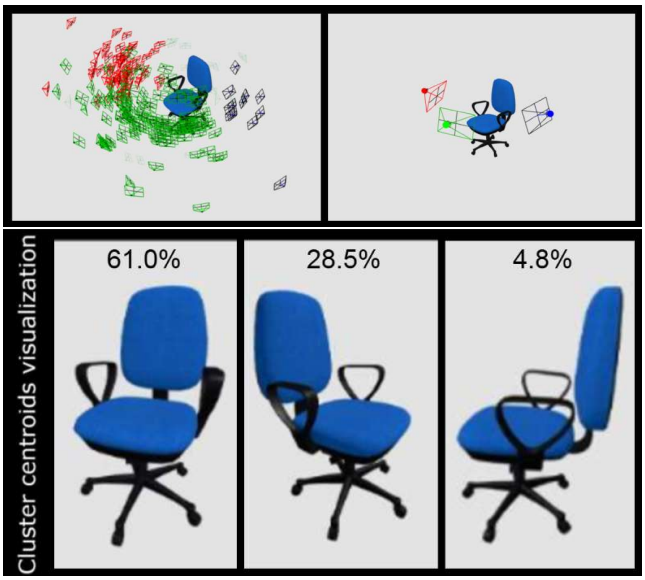

Figure 4: Spatial distribution of camera frusta and cluster identification. Top-left: the distribution of captured user POVs, color-coded by the three clusters. Top-right: the three centroids. Bottom: the view seen by these three viewpoints (with the proportion of the total captured user POVs).

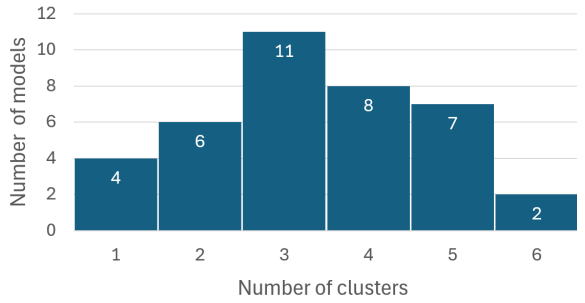

Figure 5: Frequency distribution of identified clusters across the 3D model dataset.

Clustered viewpoints. Fragmented observational patterns ($k = 4$ to 6) emerge when models possess multiple informative sides or symmetrical structures. In cases of geometric symmetry, such as the `chair` or `backpack` assets, users naturally investigate multiple equally relevant viewpoints, preventing the emergence of a single dominant perspective. However, even when a unique cluster fails to define a globally optimal viewpoint (see Fig. 7), the gathered statistics remain informative. In such instances, the clusters often align along a consistent latitude, defining a globally acceptable "elevation zone" from which the majority of users inspected the object. This suggests that while horizontal symmetry may induce fragmentation across the azimuth, user behavior remains constrained by a shared preference for specific vertical viewing angles, effectively mapping the most significant visual horizon for the asset.

Table 1: Examples of viewpoint clustering statistics for models exhibiting viewpoint consensus (top), clustered viewpoints (middle), and scattered viewpoints.

3D Model	Viewpoints	k	$\bar{d}$	$MCP(\%)$	CI
`statue`	258	1	2.12	95.0	94.96
`forniture`	265	1	1.86	95.1	95.09
`table1`	232	1	1.94	89.2	89.22
`vase2`	58	1	2.25	96.6	96.55
`chair`	232	2	1.68	64.2	32.11
`plate`	164	4	2.02	60.4	15.09
`backpack2`	277	5	1.98	30.3	6.06
`shoe`	212	6	1.76	35.8	5.97
`camera`	302	6	1.99	28.5	4.75
`camera2`	258	6	2.02	43.4	7.24
`car`	283	5	1.67	39.9	7.99
`shoe2`	236	4	1.83	47.9	11.97

Scattered viewpoints. In complex assets such as `camera`, fragmentation reflects a split interest between competing functional interfaces (e.g., lens vs. screen). This also includes cases where no clear preference emerges (see Fig. 8). This likely indicates the need for larger datasets to overcome the noise in limited observations. Nevertheless, even these low-confidence statistics remain valuable for applications like e-commerce, as they can directly generate *importance maps* to drive 3D modeling priorities or digitization focus.

5. Conclusions and road ahead

In this work, we propose a flexible framework usable for harvesting viewpoints for a given 3D model, and use it for a first capture session. We publicly provide the captured benchmark (consisting of viewpoints and complete with data on models, pertecipants - anonymized), the questionnaire used, as well as the tools to build custom questionnaires and replicate the preliminary analysis of the results, are available on the official project page:

`https://mips.di.unimi.it/pov-harvest/`

Our aspiration is that this will serve as a ground truth for assessing existing or future "best view" techniques, and, most importantly, to pave the road to a new category of data-driven approaches.

Acknowledgments

We are grateful to all the participants to our study. We thank the anonymous reviewers. This work was partially supported by the PRIN project nr.202273Z7PZ (I-CLOTH) by Italian MUR.

References

[Ame17] AMERICAN PSYCHOLOGICAL ASSOCIATION: Ethical principles of psychologists and code of conduct, 2017. 4

[BBT*13] BRIVIO P., BENEDETTI L., TARINI M., PONCHIO F., CIGNONI P., SCOPIGNO R.: Photocloud: Interactive remote exploration of joint 2d and 3d datasets. *IEEE Computer Graphics and Applications 33*, 2 (2013), 86–96, c3. 1

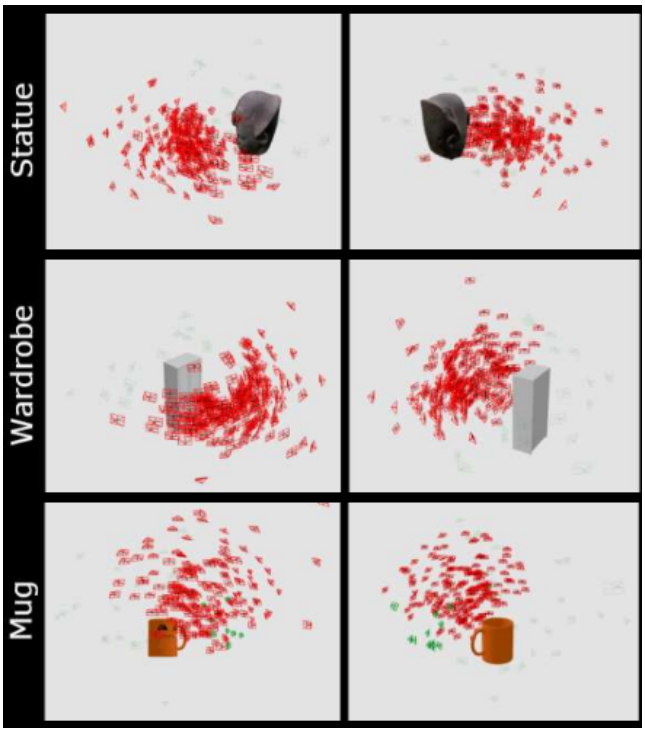

Figure 6: Examples of dataset where the harvested viewpoints exhibited consensus. In the Statue revealed a consistent preference for frontal views. For the Wardrobe, the natural view position was preferred, revealing a functional view bias. For the Mug case, views were attracted to the textured side.

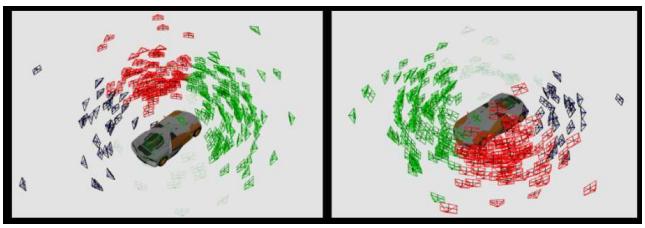

Figure 7: In this 3D model of a car, the points of view selected by the subjects are separated in three different clusters (color coded).

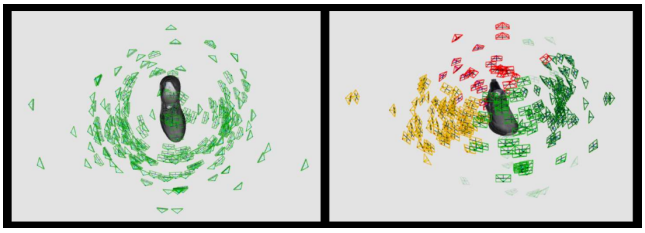

Figure 8: In this 3D model of a shoe, the selected viewpoints appear widely dispersed, showing no clear convergence or dominant viewing direction among subjects.

[BFS*18] BONAVENTURA X., FEIXAS M., SBERT M., CHUANG L., WALLRAVEN C.: A survey of viewpoint selection methods for polygonal models. *Entropy 20*, 5 (2018). 2

[BN07] BAR M., NETA M.: Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia 45*, 10 (2007), 2191–2200. 2

[BPGG16] BERTAMINI M., PALUMBO L., GHEORGHES T. N., GALATSIDAS M.: Do observers like curvature or do they dislike angularity? *British Journal of Psychology 107*, 1 (2016), 154–178. 2

[BTB99] BLANZ V., TARR M. J., BÜLTHOFF H.: What object attributes determine canonical views? *Perception 28*, 5 (1999), 575–599. 2

[DCG10] DUTAGACI H., CHEUNG C. P., GODIL A.: A benchmark for best view selection of 3D objects. In *Proc. of the ACM Workshop on 3D Object Retrieval* (NY, USA, 2010), 3DOR '10, ACM, pp. 45–50. 2

[FSG09] FEIXAS M., SBERT M., GONZÁLEZ F.: A unified information-theoretic framework for viewpoint selection and mesh saliency. *ACM Trans. Appl. Percept. 6*, 1 (Feb. 2009). 2

[HSO*22] HARTWIG S., SCHELLING M., ONZENOODT C. V., VÁZQUEZ P.-P., HERMOSILLA P., ROPINSKI T.: Learning human viewpoint preferences from sparsely annotated models. *Computer Graphics Forum 41*, 6 (2022), 453–466. 2

[JMF*25] JOOS L., MOONEY G. J., FISCHER M. T., KEIM D. A., SCHREIBER F., PURCHASE H. C., KLEIN K.: Show me your best side: Characteristics of user-preferred perspectives for 3D graph drawings. *arXiv preprint arXiv:2506.09212* (2025). 2

[Lag10] LAGA H.: Semantics-driven approach for automatic selection of best views of 3d shapes. In *3DOR@ Eurographics* (2010), pp. 15–22. 2

[LST12] LEIFMAN G., SHTROM E., TAL A.: Surface regions of interest for viewpoint selection. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012), pp. 414–421. 2

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. *ACM Trans. Graph. 24*, 3 (July 2005), 659–666. 2

[MCS16] MALOMO L., CIGNONI P., SCOPIGNO R.: Generalized trackball for surfing over surfaces. In *Proc. of Smart Tools and App. in Comp. Graph.* (2016), STAG '16, Eurographics Association, p. 89–97. 4

[MS09] MORTARA M., SPAGNUOLO M.: Semantics-driven best view of 3d shapes. *Computers & Graphics 33*, 3 (2009), 280–290. IEEE International Conference on Shape Modelling and Applications 2009. 2

[Nat79] NATIONAL COMMISSION FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH: The belmont report: Ethical principles and guidelines for the protection of human subjects of research, 1979. US Dept. Health, Edu., and Welf. 4

[NCL15] NOURI A., CHARRIER C., LÉZORAY O.: Multi-scale mesh saliency with local adaptive patches for viewpoint selection. *Signal Processing: Image Communication 38* (2015), 151–166. 2

[Nor13] NORMAN D.: *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, New York, 2013. 1

[Ske26] SKETCHFAB: Sketchfab 3D model platform. https://sketchfab.com, 2026. Accessed: 2026-02-22. 5

[SLF*11] SECORD A., LU J., FINKELSTEIN A., SINGH M., NEALEN A.: Perceptual models of viewpoint preference. *ACM Trans. Graph. 30*, 5 (Oct. 2011). 2

[SSC16] STUCCHI N., SCOCCHIA L., CARLINI A.: When geometry constrains vision: Systematic misperceptions within geometrical configurations. *PLOS ONE 11*, 3 (03 2016), 1–19. 2

[VFSH03] VÁZQUEZ P.-P., FEIXAS M., SBERT M., HEIDRICH W.: Automatic view selection using viewpoint entropy and its application to image-based modelling. *Computer Graphics Forum 22*, 4 (2003), 689–700. 2

[WEK*12] WAGEMANS J., ELDER J. H., KUBOVY M., PALMER S. E., PETERSON M. A., SINGH M., VON DER HEYDT R.: A century of gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization. *Psychological Bulletin 138*, 6 (2012), 1172–1217. 2