

Article

# A Survey of Viewpoint Selection Methods for Polygonal Models

Xavier Bonaventura <sup>1</sup> , Miquel Feixas <sup>1</sup> , Mateu Sbert <sup>1,2,\*</sup> , Lewis Chuang <sup>3</sup>  and Christian Wallraven <sup>4</sup> 

<sup>1</sup> Graphics & Imaging Laboratory, University of Girona, Girona 17003, Spain;  
xavibonaventura@gmail.com (X.B.); feixas@ima.udg.edu (M.F.)

<sup>2</sup> School of Computer Science and Technology, Tianjin University, Tianjin 300350, China

<sup>3</sup> Max Planck Institute for Biological Cybernetics, Tuebingen 72076, Germany;  
lewis.chuang@tuebingen.mpg.de

<sup>4</sup> Department of Brain Cognitive Engineering, Korea University, Seoul 02841, Korea;  
christian.wallraven@gmail.com

\* Correspondence: mateu.sbert@gmail.com

Received: 24 March 2018; Accepted: 11 May 2018; Published: 16 May 2018



**Abstract:** Viewpoint selection has been an emerging area in computer graphics for some years, and it is now getting maturity with applications in fields such as scene navigation, scientific visualization, object recognition, mesh simplification, and camera placement. In this survey, we review and compare twenty-two measures to select good views of a polygonal 3D model, classify them using an extension of the categories defined by Secord et al., and evaluate them against the Dutagaci et al. benchmark. Eleven of these measures have not been reviewed in previous surveys. Three out of the five short-listed best viewpoint measures are directly related to information. We also present in which fields the different viewpoint measures have been applied. Finally, we provide a publicly available framework where all the viewpoint selection measures are implemented and can be compared against each other.

**Keywords:** visualization; viewpoint selection; entropy; mutual information

---

## 1. Introduction

Why is viewpoint selection important? A large number of 3D models or objects are daily used across diverse fields such as computer game development, computer-aided design, and interior design. These models are often obtained by exploring large 3D model databases in as little time as possible. In this case, automated viewpoint selection plays an important role since such an application can show the model view that allows for ready recognition or understanding of the underlying 3D model. An ideal view should strive to capture the maximum information of the 3D model, such as its main characteristics, parts, functionalities, etc. The quality of this view could affect the number of models that the artist can explore in a certain period of time.

In the viewpoint selection study, the basic question is “what are good views of a 3D object or a scene?” In order to address this, a number of computational measures have been proposed to quantify the goodness or the quality of a view. Depending on our goals, the best viewpoint can be, for instance, the view that allows us to see the largest number of parts of the object, the view that shows the most salient regions of the object, or the view that maximally changes when the underlying object is jittered.

The human visual system is classically described [1] either in terms of its ability to recognize familiar three-dimensional objects as structural representations of their comprising part-components [2], or as multiple-view descriptions [3–5]. Biederman [2] proposed that familiar

object recognition can be conceptualized as a computational process by which the projected retinal image of the three-dimensional object is segmented at regions of deep concavity to derive a reduced representation of its simple geometric components (e.g., blocks, cylinders, wedges, and cones) and their spatial relations. Nonetheless, many studies have since demonstrated that the visual system demonstrates preferential behavioral and neuronal responses to particular object views [5–7]. Indeed, recognition behavior continues to be highly selective for previously learned views even when highly unique object parts with little self-occlusion are made available for discrimination [6]. Naturally, this raises the question of which view(s) could represent a given object, so as to support robust visual recognition. Palmer et al. [8] found that participants tend to agree on the canonical view (or most representative image) of each familiar object that would facilitate its recognition.

They are often off-axis views, such as a top-down three-quarter view, which arguably reveals the largest amount of surface area. In contrast, Harman et al. [9] allowed participants to learn novel 3D objects by exploring them in virtual reality. They found that their participants spent time exploring “plan” views, namely views that were on-axis or orthogonal and parallel to the object’s structural axis. Perrett et al. [10,11] found a similar preference for “plan” views in tool-like as well as “novel” objects. The mixed evidence could be due to the fact that view-canonicity can be expressed by the multiple factors [12]: goodness for recognition (a good view for recognition shows the most salient and significant features, and it is stable with respect to small transformations, and it avoids a high number of occluded features), familiarity (recognition is influenced by the views that are encountered more frequently and during the initial learning), functionality (recognition is influenced by the views that are most relevant for how we interact with an object), and aesthetic criteria (preferred views can be influenced by geometric proportions).

In this survey, the computational measures that will be reviewed are those that were motivated for “goodness for recognition” instead of other aspects such as familiarity and aesthetics. The main contribution of this survey lies in collecting and testing in a common framework the most basic measures introduced to select the best views for polygonal models. We include most of the measures presented in previous surveys of viewpoint selection [13–15], but we do not consider semantic-based viewpoint selection measures, the absolute Gaussian and mean curvature [15], and the topological complexity [13]. In addition, we review eleven viewpoint selection measures that have not been included in the previous surveys.

This survey is organized as follows. In Section 2, we review pioneering work in view-selection and the basic measures that have been proposed for estimating the quality of views. In Section 3, the most relevant measures are defined and described. In Section 4, we test the presented measures using the Dutagaci et al. [14] benchmark. In Section 5, we present literature that applies the viewpoint quality measures to other fields of research. Finally, in Section 6, our conclusions and future work are presented.

## 2. Background

In this section, we present the basis of viewpoint selection, that is, landmark research and the most basic measures that gave rise to the other measures and methods that have also been used in the last decade (Section 3).

First, we review pioneer work on viewpoint selection. Attneave [16] analyzes informational aspects of visual perception and explains that information for object discrimination is concentrated along an object’s contour shape (i.e., 2D silhouette), especially where such information changes rapidly (i.e., peaks of curvature). Connolly [17] describes two algorithms that use partial octree models to determine the next best view to take. Kamada and Kawai [18] presented a measure to select a good view based on the angle between the view direction and the normal of the planes of the model. This method tries to avoid degenerative views, views where a plane is projected as a line and a line is projected as a point. Plemenos and Benayada [19] extended Kamada’s work to ensure that the user sees a great number of details. Plemenos’ measure takes into account the projected area and the number

of polygons to evaluate the viewpoint goodness. Arbel and Ferrie [20] applied Shannon entropy to define entropy maps to guide an active observer along an optimal trajectory. Inspired by Kamada's and Plemenos' works, Vázquez et al. [21] also used the Shannon entropy to quantify the information provided by a view. This measure incorporates both the projected area and the number of faces.

Weinshall and Werman [22] define two measures: view likelihood and view stability. View likelihood measure is used to identify "characteristic" views based on the probability that a certain view of a given 3D object is observed. View stability is used to identify "generic" views based on how the image changes as the viewpoint is slightly modified. Stoev and Straßer [23] noticed that the projected area was not enough to visualize terrains and they presented a method that maximizes the maximum depth of the image in addition to the projected area. Given a sphere of viewpoints, Yamauchi et al. [24] computed the similarity between each two disjoint views using Zernike moments analysis and obtained a similarity weighted spherical graph. Here, a view was considered to be stable if all of the edges that were incident on its viewpoint in the spherical graph had high similarity weights.

Itti et al. [25] maintain that visual attention is saliency-dependent and use a saliency map to represent the conspicuity or saliency at every location in the visual field by a scalar quantity. Thus, a good view could be described as one that is likely to be attended to, given its high saliency content. Borji and Itti [26] presented a state-of-the-art in visual attention modeling that can compute saliency maps from any image or video input. From surface curvature, Lee et al. [27] introduced a perception-inspired measure of regional importance, called mesh saliency, that has been used in mesh simplification and viewpoint selection. Gal and Cohen-Or [28] introduced a method for partial matching of surfaces by using the abstraction of salient geometric features and a method to construct them.

Some measures that consider semantic information of the model have been also used in viewpoint selection. High-level and semantic measures take into account features such as the topology of the model, the position of the eyes, or the part used to grasp the object. Becker et al. [29] analyze how object-intrinsic oddities can be detected by previous semantic knowledge of the object, and that draw the attention of the viewer by its oddity. Koulieris et al. [30] define a high-level saliency model for objects within a scene, based on singleness and semantic coherence with environment objects, that allow to identify objects to be rendered in a higher detail. To include this saliency model in the view point selection process, a priori semantic information about the objects constituting the scene is needed. Secord et al. [15], based on the work of Blanz et al. [12] and Gooch et al. [31], propose a measure that captures views from slightly above the horizon. Secord et al. [15] also introduced a measure that tends to avoid views from directly below for objects that have an obvious orientation. The automatic method of Fu et al. [32] can be used to determine both the base and the orientation of the object. When the model is a creature with eyes or a face, people prefer views where the eyes are visible [33]. Secord et al. [15] have further proposed an attribute that sums all the visible pixels corresponding to the eyes' surface. Finally, it is worth mentioning that Podolak et al. [34] have introduced a method to choose good viewpoints automatically by minimizing the symmetry of the object seen from the viewpoint.

Polonsky et al. [13] and Secord et al. [15] have described and analyzed a number of measures that were introduced to quantify the goodness of a view of an object. After analyzing different view descriptors, Polonsky et al. [13] concluded that no single descriptor does a perfect job and have suggested that a combination of descriptors would amplify their respective advantage over each other. In this regard, Secord et al. [15] have presented a perceptual model of viewpoint selection based on the combination of different attributes such as surface visibility, silhouette length, projected area, and maximum depth. If the region corresponding to the eyes' surface is marked, Secord et al. [15] have proposed changing the maximum depth according to eye preference.

Dugataci et al. [14] have presented a benchmark to validate best view selection methods by analyzing the accuracy of these methods in comparison with the preferred views selected by 26 human subjects. In this benchmark, the human subjects were asked to select the most informative view of 68 3D models through a web page. Dugataci et al. [14] also compute for every model the inconsistency

of the choices of the human subjects. They provide a way to quantify the error of a best view selection algorithm compared to the data collected. An error between 0 and 1 and the average for all the models can be computed using the benchmark. To compute the error, they take into account the symmetry of the models. Most of the models used in this benchmark are common objects highly familiar to humans. The benchmark was tested with seven different methods computed in a sphere of 258 viewpoints. The methods tested were view area, ratio of visible area, surface area entropy, silhouette length, silhouette entropy, curvature entropy, and mesh saliency.

### 3. Viewpoint Selection Measures

In this section, we gather twenty-two viewpoint selection measures that are classified according to several attributes captured from a particular viewpoint: area, silhouette, depth, stability, and surface curvature. These categories, except stability, are presented in Secord et al. [15]. For each measure, we provide its definition and the reference of the paper where the measure was introduced. All the measures presented in this section will be tested in Section 4 and are available in a public common framework.

#### 3.1. Notation

For comparison purposes, we propose a unified notation for the analyzed measures adopted from Feixas et al. [35], where an information channel was defined between a set of viewpoints  $\mathcal{V}$  and a set of polygons  $\mathcal{Z}$ . The projected area of polygon  $z$  from viewpoint  $v$  is denoted by  $a_z(v)$  and the projected area of the model from viewpoint  $v$  is given by  $a_t(v)$ . Viewpoint quality  $v$  is expressed by  $VQ(v)$ .

In Feixas et al. [35], a viewpoint selection framework was proposed from an information channel  $V \rightarrow Z$  between the random variables  $V$  (input) and  $Z$  (output), which represent, respectively, a set of viewpoints  $\mathcal{V}$  and the set of polygons  $\mathcal{Z}$  of an object. This channel is defined by a conditional probability matrix obtained from the projected areas of polygons at each viewpoint and can be interpreted as a visibility channel where the conditional probabilities represent the probability of seeing a determined polygon from a given viewpoint. The three basic elements of the visibility channel are:

- Conditional probability matrix  $p(Z|V)$ , where each element  $p(z|v) = \frac{a_z(v)}{a_t(v)}$  is defined by the normalized projected area of polygon  $z$  over the sphere of directions centered at viewpoint  $v$ .  
Conditional probabilities fulfill  $\sum_{z \in \mathcal{Z}} p(z|v) = 1$ .
- Input distribution  $p(V)$ , where each element  $p(v) = \frac{a_t(v)}{\sum_{v \in \mathcal{V}} a_t(v)}$ , which represents the probability of selecting each viewpoint, is obtained from the normalization of the object projected area at each viewpoint. The input distribution is interpreted as the importance assigned to each viewpoint  $v$ .
- Output distribution  $p(Z)$ , given by  $p(z) = \sum_{v \in \mathcal{V}} p(v)p(z|v)$ , which represents the average projected area of polygon  $z$ .

Tables 1 and 2 show, respectively, the notation used in the measure definitions and the list of measures studied in this paper. Observe that Table 2 also contains additional information for each measure. Columns 3, 4, and 5 show the corresponding names used in surveys by Polonsky et al. [13], Dugataci et al. [14], and Secord et al. [15], respectively. Column 6 indicates whether the best viewpoint corresponds to the highest (H) or the lowest (L) measure value. Column 7 shows whether the measure is sensitive (Y) to how the polygonal model is discretized or not (N). In addition, column 8 gives the main reference of the measure presented.

**Table 1.** The most relevant notation symbols used in this paper.

$z$	polygon
$\mathcal{Z}$	set of polygons
$v$	viewpoint
$\mathcal{V}$	set of viewpoints
$a_z(v)$	projected area of polygon $z$ from viewpoint $v$
$a_t(v)$	projected area of the model from viewpoint $v$
$vis_z(v)$	visibility of polygon $z$ from viewpoint $v$ (0 or 1)
$N_p$	number of polygons
$R$	number of pixels of the projected image
$A_z$	area of polygon $z$
$A_t$	total area of the model
$p(z v)$	conditional probability of $z$ given $v$
$p(z)$	probability of $z$
$p(v z)$	conditional probability of $v$ given $z$
$p(v)$	probability of $v$
$H(V)$	entropy of the set of viewpoints
$H(Z)$	entropy of the set of polygons
$H(V z)$	conditional entropy of the set of viewpoints given polygon $z$
$H(Z v)$	conditional entropy of the set of polygons given viewpoint $v$
$strength(v)$	silhouette length from viewpoint $v$
$\{h(\alpha)\}$	normalized silhouette curvature histogram
$\alpha$	turning angle bin
$a$	turning angle between two consecutive pixels
$\mathcal{A}$	set of turning angles
$N_a$	number of turning angles
$depth(v)$	normalized maximum depth of the scene from viewpoint $v$
$\{h(d)\}$	normalized histogram of depths
$d$	depth bin
$\mathcal{D}$	set of depth bins
$N_v$	number of neighbors of $v$
$L(v)$	size of the compression of the depth image corresponding to viewpoint $v$
$L(v_i, v_j)$	size of the compression of the concatenation of the depth images corresponding to viewpoints $v_i$ and $v_j$
$K_i$	curvature of vertex $i$
$\{h(b)\}$	normalized histogram of visible curvatures from viewpoint $v$
$b$	curvature bin
$\mathcal{B}$	set of curvature bins
$S(x)$	salinity of vertex $x$

### 3.2. Area Attributes

The measures based on these attributes are computed using as main feature the area of polygons seen from a particular viewpoint.

**Number of visible triangles.** Plemenos and Benayada [19] used the number of visible triangles seen from a viewpoint as a viewpoint quality measure. The higher the number of visible triangles, the better the quality of a viewpoint. This measure is based on the fact that the most significant regions contain more details and, thus, more triangles. This measure is expressed as

$$VQ_1(v) = \sum_{z \in \mathcal{Z}} vis_z(v), \quad (1)$$

where  $vis_z(v)$  is 1 if the polygon  $z$  is visible from viewpoint  $v$  and 0 otherwise. Different criteria can be used to consider whether a polygon is visible. In our implementation, a polygon is considered visible if at least one pixel of polygon  $z$  is visible from viewpoint  $v$  ( $a_z(v) > 0$ ). Obviously, the number of visible triangles is sensitive to the discretization of the model.

**Projected area.** Plemenos and Benayada [19] also studied the projected area of the model from a viewpoint as a measure of viewpoint goodness since the number of visible triangles was found not to be enough in some cases. For example, if we consider a pencil, it is normal to have a high number of polygons around the pencil point. If we use the number of visible triangles to select the best viewpoint, we would only see a small part of the object. The projected area expressed as

$$VQ_2(v) = a_t(v) \quad (2)$$

can be considered as a viewpoint quality measure. Thus, the higher the projected area, the better the viewpoint quality. This measure is insensitive to the discretization of the model.

**Plemenos and Benayada.** Plemenos and Benayada [19] combined the number of visible triangles and the projected area to create a measure for viewpoint quality. A viewpoint is considered good if the percentage of the number of visible polygons plus the percentage of projected area with respect to the size of the screen is high. This measure can be expressed as

$$VQ_3(v) = \frac{\sum_{z \in \mathcal{Z}} \lceil \frac{a_z(v)}{a_z(v)+1} \rceil}{N} + \frac{\sum_{z \in \mathcal{Z}} a_z(v)}{R}, \quad (3)$$

where  $R$  is the total number of pixels of the image and  $N$  the total number of polygons (i.e.,  $N = |\mathcal{Z}|$ ). For more details, see also Barral et al. [36]. Note that the first term is the ratio of visible polygons, where  $\lceil \frac{a_z(v)}{a_z(v)+1} \rceil$  is equivalent to  $vis_z(v)$ , and the second term is the ratio of the projected area with respect to the resolution of the screen. Thus,  $VQ_3(v)$  can be rewritten as

$$VQ_3(v) = \frac{VQ_1(v)}{N} + \frac{VQ_2(v)}{R}. \quad (4)$$

This measure is sensitive to polygonal discretization because  $VQ_1(v)$  is, as we have seen above.

**Visibility ratio.** Plemenos and Benayada [19] also introduced the ratio between the visible surface area of the model from viewpoint  $v$  and the total surface area as a viewpoint quality measure. The visibility ratio is expressed by

$$VQ_4(v) = \frac{\sum_{z \in \mathcal{Z}} vis_z(v) A_z}{A_t}, \quad (5)$$

where  $A_z$  is the area of polygon  $z$ , and  $A_t$  is the total area of the model. Observe that  $A_z$  does not depend on the viewpoint because denotes the real area of polygon  $z$ . The best viewpoint corresponds to the minimum value of the measure. This measure is insensitive to the discretization of the model.

**Viewpoint entropy.** Vázquez et al. [21,37] presented a measure for viewpoint selection based on Shannon entropy [38,39]. This measure takes into account the projected area and the number of viewpoints and can be understood as the amount of information captured by a specific viewpoint. The viewpoint entropy is defined by

$$VQ_5(v) = H(v) = - \sum_{z \in \mathcal{Z}} \frac{a_z(v)}{a_t(v)} \log \frac{a_z(v)}{a_t(v)}. \quad (6)$$

Using the notation of the visibility channel introduced in Section 3.1, the viewpoint entropy is rewritten as

$$VQ_5(v) = H(Z|v) = - \sum_{z \in \mathcal{Z}} p(z|v) \log p(z|v), \quad (7)$$

where  $H(Z|v)$  represents the conditional entropy of  $Z$  given a viewpoint  $v$ . The best viewpoint corresponds to the one with maximum entropy, which is obtained when a certain viewpoint can see all the faces with the same relative projected area. Viewpoint entropy is sensitive to polygonal discretization as in general the entropy increases with the number of polygons.

Polonsky et al. [13] propose the application of viewpoint entropy using the probability of semantically important segments of the model.

**Information  $I_2$ .** Deweese and Meister [40] used a decomposition of mutual information in the field of neuroscience to quantify the information associated with stimuli and responses. Bonaventura et al. [41] applied this measure to the field of best viewpoint selection to express the informativeness of a viewpoint. The viewpoint information  $I_2$  is defined by

$$\begin{aligned}
VQ_6(v) = I_2(v; Z) &= H(Z) - H(Z|v) \\
&= H(Z) - VQ_5(v) \\
&= - \sum_{z \in \mathcal{Z}} p(z) \log p(z) + \sum_{z \in \mathcal{Z}} p(z|v) \log p(z|v),
\end{aligned} \tag{8}$$

where  $H(Z)$  stands for the entropy of model triangles. Note that  $I_2$  is closely related to viewpoint entropy, defined as  $H(Z|v)$  [21,35], since  $I_2(v; Z) = H(Z) - H(Z|v)$ . As  $H(Z)$  is constant for a given mesh resolution,  $I_2(v; Z)$  and viewpoint entropy have the same behavior in viewpoint selection because the highest value of  $I_2(v; Z)$  corresponds to the lowest value of viewpoint entropy, and vice versa. An important drawback of viewpoint entropy is that it goes to infinity for finer and finer resolutions of the mesh [35], while  $I_2$  presents a more stable behavior due to the normalizing effect of  $H(Z)$  in Equation (8). The best viewpoint is given by the one that has minimum  $I_2$ . Similarly to viewpoint entropy, this measure is also sensitive to polygonal discretization.

**Viewpoint Kullback–Leibler distance (VKL).** Sbert et al. [42] presented a viewpoint quality measure given by the Kullback–Leibler distance between the normalized distribution of the projected areas of polygons from viewpoint  $v$  and the normalized distribution of the real areas of polygons. The viewpoint Kullback–Leibler distance is given by

$$VQ_7(v) = \sum_{z \in \mathcal{Z}} \frac{a_z(v)}{a_t(v)} \log \frac{\frac{a_z(v)}{a_t(v)}}{\frac{A_z}{A_t}}. \tag{9}$$

Observe that the minimum value, which corresponds to the best viewpoint, is obtained when the normalized distribution of projected areas is equal to the normalized distribution of real areas. Viewpoint Kullback–Leibler distance is near insensitive to polygonal discretization.

**Viewpoint mutual information (or  $I_1$ ).** Feixas et al. [35] presented a measure, called viewpoint mutual information (VMI), that captures the degree of correlation between a viewpoint and the set of polygons. Bonaventura et al. [41] renamed this measure as  $I_1$  because this is one of the decomposition forms of mutual information used to deal with stimuli and responses [40]. The viewpoint mutual information is defined by

$$VQ_8(v) = VMI(v) = I_1(v; Z) = \sum_{z \in \mathcal{Z}} p(z|v) \log \frac{p(z|v)}{p(z)}. \tag{10}$$

High values of the measure mean a high correlation between viewpoint  $v$  and the object, indicating a highly coupled view (for instance, between the viewpoint and a small number of polygons with low average visibility). On the other hand, the lowest values correspond to the most representative or relevant views (i.e., best viewpoints), showing the maximum possible number of polygons in a balanced way. VMI is insensitive to the discretization of the model. For more information, see [43].

**Information  $I_3$ .** Butts [44] introduced a new decomposition form of mutual information, called  $I_3$ , to quantify the specific information associated with a stimulus. Bonaventura et al. [41] proposed  $I_3$  as a viewpoint quality measure. The measure  $I_3$  is defined by

$$VQ_9(v) = I_3(v; Z) = \sum_{z \in \mathcal{Z}} p(z|v) I_2(V; z), \tag{11}$$

where  $I_2(V; z)$  is the specific information of polygon  $z$  given by

$$\begin{aligned}
I_2(V; z) &= H(V) - H(V|z) \\
&= - \sum_{v \in \mathcal{V}} p(v) \log p(v) + \sum_{v \in \mathcal{V}} p(v|z) \log p(v|z),
\end{aligned} \tag{12}$$

where  $p(v|z) = \frac{p(v)p(z|v)}{p(z)}$  (Bayes theorem). Note that  $H(V)$  and  $H(V|z)$  represent the entropy of the set of viewpoints and the conditional entropy of the set of viewpoints given polygon  $z$ , respectively. A high value of  $I_3(v; Z)$  means that the polygons seen by  $v$  are very informative in the sense of  $I_2(V; z)$ . The most informative viewpoints are considered as the best views and correspond to the viewpoints that see the highest number of maximally informative polygons. The measure  $I_3$  is sensitive to polygonal discretization.

### 3.3. Silhouette Attributes

The measures based on these attributes are computed using the silhouette of the object seen from a particular viewpoint. All these measures are insensitive to the discretization of the model because the polygons are not directly used.

**Silhouette length.** Polonsky et al. [13] presented the silhouette length of the projected model from a viewpoint  $v$  as a measure of viewpoint goodness. The silhouette length is expressed as

$$VQ_{10}(v) = \text{slength}(v), \quad (13)$$

where  $\text{slength}(v)$  stands for the silhouette length from  $v$ . In our implementation, the silhouette length of the model is computed from the viewpoint  $v$  by counting the number of pixels that belong to the silhouette. If there are multiple contours, the pixels of all the contours are added. The goodness of a viewpoint is associated with the maximum silhouette length.

**Silhouette entropy.** Polonsky et al. [13] introduced the entropy of the silhouette curvature distribution, proposed by Page et al. [45], as a measure of viewpoint goodness. In our implementation, the silhouette curvature histogram is computed from the turning angles between consecutive pixels belonging to the silhouette. The range of the curvature is between  $-\pi/2$  and  $\pi/2$  with a step of  $\pi/4$  due to the angles obtained between neighbor pixels. The silhouette entropy is defined by

$$VQ_{11}(v) = - \sum_{\alpha=-\pi/2}^{\pi/2} h(\alpha) \log h(\alpha), \quad (14)$$

where  $\{h(\alpha)\}$  represents the normalized silhouette curvature histogram and  $\alpha$  is the turning angle bin. The best viewpoint is the one with the highest silhouette entropy.

**Silhouette curvature.** Vieira et al. [46] introduced the complexity of the silhouette defined as the total integral of its curvature. In our implementation, the silhouette curvature is computed as

$$VQ_{12}(v) = \frac{\sum_{c \in \mathcal{C}} \frac{|c|}{\frac{\pi}{2}}}{N_c}, \quad (15)$$

where  $c$  is the turning angle between two consecutive pixels,  $\mathcal{C}$  is the set of turning angles, and  $N_c$  is the number of turning angles, equal to the number of pixels of the silhouette. The best viewpoint is given by the one with the maximum value.

**Silhouette curvature extrema.** As a variation of the above silhouette curvature measure, Secord et al. [15] introduced the silhouette curvature extrema to emphasize high curvatures on the silhouette. The silhouette curvature extrema is computed as

$$VQ_{13}(v) = \frac{\sum_{c \in \mathcal{C}} \left( \frac{c}{\frac{\pi}{2}} \right)^2}{N_c}. \quad (16)$$

Similarly to silhouette curvature, the higher the value, the better the viewpoint.

### 3.4. Depth Attributes

The measures based on these attributes are computed using the depth of the model seen from a particular viewpoint.

**Stoev and Straßer.** Stoev and Straßer [23] noticed that the projected area was not enough to visualize terrains because usually the view with most projected area is the one from above. They presented a method for camera placement that maximizes the maximum depth of the image in addition to the projected area. This measure is defined by

$$VQ_{14}(v) = \alpha p(v) + \beta d(v) + \gamma(1 - |d(v) - p(v)|), \quad (17)$$

where  $p(v)$  is the normalized projection area from viewpoint  $v$  and  $d(v)$  is the normalized maximum depth of the scene from viewpoint  $v$ . For general purposes, the authors proposed the use of the following values:  $\alpha = \beta = \gamma = \frac{1}{3}$ . The Stoev and Straßer measure used in our implementation is given by

$$VQ_{14}(v) = \frac{1}{3}p(v) + \frac{1}{3}d(v) + \frac{1}{3}(1 - |d(v) - p(v)|). \quad (18)$$

For terrain scenarios, Stoev and Straßer [23] considered  $\alpha = \beta = \frac{1}{4}$  and  $\gamma = \frac{1}{2}$ . The best viewpoint is the one with the maximum value, maximizing the projected area and the maximum depth and minimizing the difference between the projected area and the maximum depth. This measure is insensitive to polygonal discretization because the projected area and the maximum depth are insensitive too.

**Maximum depth.** Secord et al. [15] considered only the maximum depth, used in Stoev and Straßer [23], as a descriptor of viewpoint quality. This measure is thus defined as

$$VQ_{15}(v) = \text{depth}(v), \quad (19)$$

where  $\text{depth}(v)$  is the maximum depth. As we have seen above, the maximum depth is insensitive to polygonal discretization and the best viewpoint is considered as the one with the maximum value.

**Depth distribution.** Instead of using only the maximum depth from a viewpoint, Secord et al. [15] proposed a measure that maximizes the visible range of depths. The depth distribution measure defined by

$$VQ_{16}(v) = 1 - \sum_{d \in \mathcal{D}} h(d)^2 \quad (20)$$

tries to capture the maximum diversity of depths, where  $d$  represents a depth bin,  $\mathcal{D}$  is the set of depth bins, and  $\{h(d)\}$  the normalized histogram of depths. The best viewpoint corresponds to the maximum value of the measure. This measure is insensitive to the discretization of the model.

### 3.5. Stability Attributes

The measures based on these attributes compute the stability of a viewpoint by comparing the viewpoint with its neighbors.

**Instability.** Feixas et al. [35] defined viewpoint instability from the notion of dissimilarity between two viewpoints, which is given by the Jensen–Shannon divergence [47] between their respective projected area distributions. The use of Jensen–Shannon as a measure of view similarity was proposed by Bordoloi and Shen [48] in the volume rendering field. The viewpoint instability of  $v$  is defined by

$$VQ_{17}(v) = \frac{1}{N_v} \sum_{j=1}^{N_v} D(v, v_j), \quad (21)$$

where  $v_j$  is a neighbor of  $v$ ,  $N_v$  is the number of neighbors of  $v$ , and

$$D(v, v_j) = JS \left( \frac{p(v)}{p(v) + p(v_j)}, \frac{p(v_j)}{p(v) + p(v_j)}; p(Z|v), p(Z|v_j) \right)$$

is the Jensen–Shannon divergence between the distributions  $p(Z|v)$  and  $p(Z|v_j)$  captured by  $v$  and  $v_j$  with weights  $\frac{p(v)}{p(v) + p(v_j)}$  and  $\frac{p(v_j)}{p(v) + p(v_j)}$ , respectively. The best viewpoint is the one with the lowest instability. The instability measure is sensitive to the discretization of the model.

**Depth-based visual stability.** Vázquez [49] introduced a method to compute the view stability from the depth images of all viewpoints. The degree of similarity between two viewpoints is given by the normalized compression distance (NCD) between two depth images:

$$\text{similarity}(v_i, v_j) = NCD(v_i, v_j) = \frac{L(v_i v_j) - \min \{L(v_i), L(v_j)\}}{\max \{L(v_i), L(v_j)\}}, \quad (22)$$

where  $L(v_i)$  and  $L(v_j)$  are, respectively, the sizes of the compression of the depth images corresponding to viewpoints  $v_i$  and  $v_j$ , and  $L(v_i v_j)$  is the size of the compression of the concatenation of the depth images corresponding to  $v_i$  and  $v_j$ .

Two views are considered similar if their distance is less than a given threshold. Hence, the most stable view is given by the one that has the largest number of similar views. The depth-based visual stability is given by

$$VQ_{18}(v) = \#\text{similar views to } v. \quad (23)$$

This measure is robust to the discretization of the model because an image-based method is used. However, it is highly sensitive to the threshold value. The best view corresponds to the most stable one.

### 3.6. Surface Curvature Attributes

The measures based on these attributes are computed using the surface curvature of the shape. Note that, in the last two measures (Equations (28) and (29)), area attributes are also taken into account.

**Curvature entropy.** Polonsky et al. [13] propose a measure that evaluates the entropy of the curvature distribution over the visible portion of surface from a given viewpoint. This measure is inspired by the entropy of the Gaussian curvature distribution defined by Page et al. [45]. The curvature of vertex  $i$  is defined by

$$K_i = 2\pi - \sum_j \phi_j, \quad (24)$$

where the angle  $\phi_j$  is the wedge subtended by the edges of a triangle whose corner is at the vertex  $i$ . The curvature entropy of a viewpoint  $v$  is defined by

$$VQ_{19}(v) = - \sum_{b \in \mathcal{B}} h(b) \log h(b), \quad (25)$$

where  $b$  represents a curvature bin,  $\mathcal{B}$  is the set of curvature bins, and  $\{h(b)\}$  the normalized histogram of visible curvatures from viewpoint  $v$ . The higher the value, the better the viewpoint. Curvature entropy is sensitive to the discretization of the model.

**Table 2.** List of measures (columns 1 and 2) grouped into five different categories with the corresponding names (columns 3, 4, and 5) used in surveys of Polonsky et al. [13], Dutagaci et al. [14], and Secord et al. [15], respectively. Column 6 indicates whether the best viewpoint corresponds to the highest (H) or the lowest (L) measure value. Column 7 shows whether the measure is sensitive (Y) to the polygonal discretization or not (N). Column 8 gives the main reference of the measure. Note: Character ‘-’ indicates that the measure was not tested in the corresponding survey.

#	Measure	Polonsky05	Dutagaci10	Secord11	V	D	Ref.
1	# Visible triangles	-	-	-	H	Y	[19]
2	Projected area	-	View area	<i>Idem</i>	H	N	[19]
3	Plemenos and Benayada	-	-	-	H	Y	[19]
4	Visibility ratio	<i>Idem</i>	Ratio of visible area	Surface visibility	H	N	[19]
5	Viewpoint entropy	Surface area entropy	Surface area entropy	<i>Idem</i>	H	Y	[21]
6	$I_2$	-	-	-	L	Y	[41]
7	VKL	-	-	-	L	N	[42]
8	VMI ( $I_1$ )	-	-	-	L	N	[35]
9	$I_3$	-	-	-	H	Y	[41]
10	Silhouette length	<i>Idem</i>	<i>Idem</i>	<i>Idem</i>	H	N	[13]
11	Silhouette entropy	<i>Idem</i>	<i>Idem</i>	-	H	N	[13]
12	Silhouette curvature	-	-	<i>Idem</i>	H	N	[46]
13	Silhouette curvature extrema	-	-	<i>Idem</i>	H	N	[15]
14	Stoev and Straßer	-	-	-	H	N	[23]
15	Maximum depth	-	-	Max depth	H	N	[23]
16	Depth distribution	-	-	<i>Idem</i>	H	N	[15]
17	Instability	-	-	-	L	Y	[35]
18	Depth-based visual stability	-	-	-	H	N	[49]
19	Curvature entropy	<i>Idem</i>	<i>Idem</i>	-	H	Y	[13]
20	Visible saliency	-	Mesh saliency	Mesh saliency	H	Y	[27]
21	Projected saliency	-	-	-	H	Y	[35]
22	Saliency-based EVMI	-	-	-	L	Y	[35]

**Visible saliency.** Lee et al. [27] presented a measure to select the best viewpoint based on the amount of saliency seen from a viewpoint. The saliency used is presented by Lee et al. [27] and it is computed for every vertex using the curvature presented by Taubin [50]. The visible saliency measure is the sum of all the saliences of the vertices seen from viewpoint  $v$  and is defined by

$$VQ_{20}(v) = \sum_{x \in \mathcal{X}} S(x), \quad (26)$$

where  $\mathcal{X}$  is the set of visible vertices and  $S(x)$  the saliency of vertex  $x$ . The saliency of vertex  $x$  is defined by

$$S(x) = |G(C(x), \sigma) - G(C(x), 2\sigma)|, \quad (27)$$

where  $G(C(v), \sigma)$  is the Gaussian-weighted average of the mean curvature. The higher the value, the better the viewpoint. Visual saliency is sensitive to polygonal discretization since the summation is done for the visible vertices. Similarly to Lee et al. [27], Sokolov and Plemenos [51] present a viewpoint quality measure given the sum of curvatures captured by a viewpoint where the curvature is computed as in Equation (24).

**Projected saliency.** Inspired by the visual saliency [27], Feixas et al. [35] presented a method to select the best view using the saliency of the polygons. This saliency is computed for every polygon using an information channel between polygons and viewpoints. The projected saliency is defined by

$$VQ_{21}(v) = \sum_{z \in Z} S(z)p(v|z), \quad (28)$$

where  $S(z)$  is saliency of polygon  $z$  computed as

$$S(z) = \frac{1}{N_z} \sum_{j=1}^{N_z} D(z, z_j),$$

where polygon  $z_j$  is a neighbor of polygon  $z$ ,  $N_z$  is the number of neighbors of  $z$ , and

$$D(z, z_j) = JS \left( \frac{p(z)}{p(z) + p(z_j)}, \frac{p(z_j)}{p(z) + p(z_j)}; p(V|z), p(V|z_j) \right)$$

is the Jensen–Shannon divergence between the distributions  $p(V|z)$  and  $p(V|z_j)$  with weights  $\frac{p(z)}{p(z) + p(z_j)}$  and  $\frac{p(z_j)}{p(z) + p(z_j)}$ , respectively. The higher the value the better the viewpoint. The projected saliency is sensitive to the discretization of the model. Similarly, other polygonal information measures have been projected to the viewpoints to select a good view [52].

**Saliency-based EVMi.** Feixas et al. [35] presented an extended version of viewpoint mutual information (EVMi) where the target distribution is weighted by an importance factor. The importance-based EVMi is defined by

$$VQ_{22}(v) = \sum_{z \in Z} p(z|v) \log \frac{p(z|v)}{p'(z)}, \quad (29)$$

where  $p'(z)$  is given by

$$p'(z) = \frac{p(z)i(z)}{\sum_{z \in Z} p(z)i(z)}, \quad (30)$$

where  $i(z)$  is the importance of polygon  $z$ . The saliency-based EVMi is obtained when  $i(z) = S(z)$  [35]. Similarly to VMI, the best viewpoint corresponds to the minimum value. Saliency-based EVMi is sensitive to polygonal discretization because the saliency of a polygon is sensitive too. Serin et al. [53]

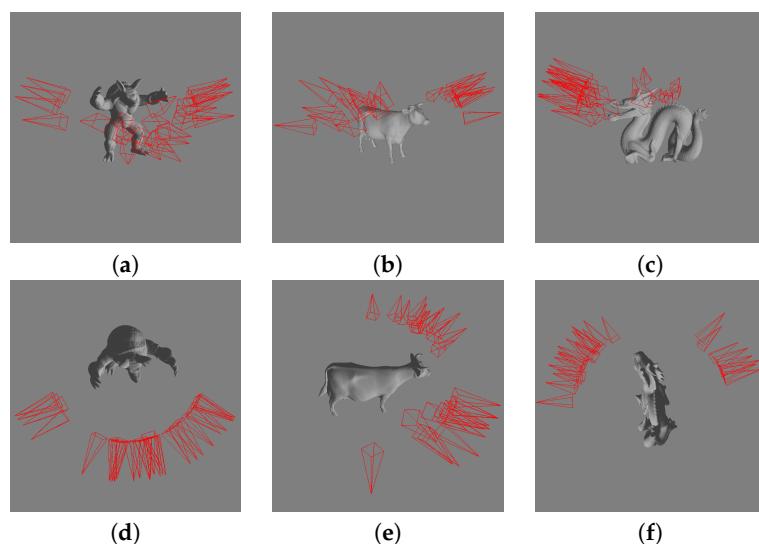
presented a similar measure where  $i(z)$  is given by the surface curvature and  $p(z)$  (i.e., average projected area) is substituted by the total area of the polygon.

#### 4. Results and Discussion

In this section, we test and compare the measures presented in Sections 3.2–3.6. These measures are computed for every model without considering any semantic information, such as the object's preferred orientation. First, we describe the details of the implementation used to compute the viewpoint selection measures. Second, we illustrate for all the measures the best view of three different 3D models. Third, the Dutagaci et al. [14] benchmark is used to analyze the accuracy of these measures in comparison with the best views selected by 26 human subjects. The presented measures, except the visual saliency measure, have been implemented in a common framework. For the visual saliency measure ( $VQ_{20}$ ), we have used Dutagaci's implementation [14]. This is the only measure not included in the framework.

To compute the projected area of a polygon (usually a triangle), we use a projection resolution of  $640 \times 640$  pixels. No back-face culling optimization is applied and the polygons are rendered from both sides. All of the models are centered inside a sphere of 642 viewpoints built from the recursive discretization of an icosahedron, and the camera is looking at the center of this sphere. The radius of the viewpoint sphere is six times the radius of the smallest bounding sphere of the model, the perspective distortion being acceptable. The view-frustum of the camera ( $19.2^\circ$ ) is adjusted to ensure that only the model and the minimum background is seen. For the results of the depth-based visual stability measure ( $VQ_{18}$ ), we use a projection resolution of  $128 \times 128$  pixels to reduce the computation time. In this case, the threshold used to decide if two viewpoints are similar is 0.87. Our framework, including the source code, is available at [54]. In this framework, the user can add and test new measures.

To show the goodness of the viewpoint quality measures, three 3D models of the Dutagaci benchmark are used: the Standford Armadillo (17,296 triangles), a cow (23,216 triangles), and the Standford dragon (26,142 triangles). Figure 1 shows the best views selected by 26 human subjects in the Dutagaci et al. [14] benchmark. Note that viewpoint entropy and information  $I_2$  are grouped in Figure 2 and in the following reported results since they have the same performance (see Equation (8) in Section 3.2).

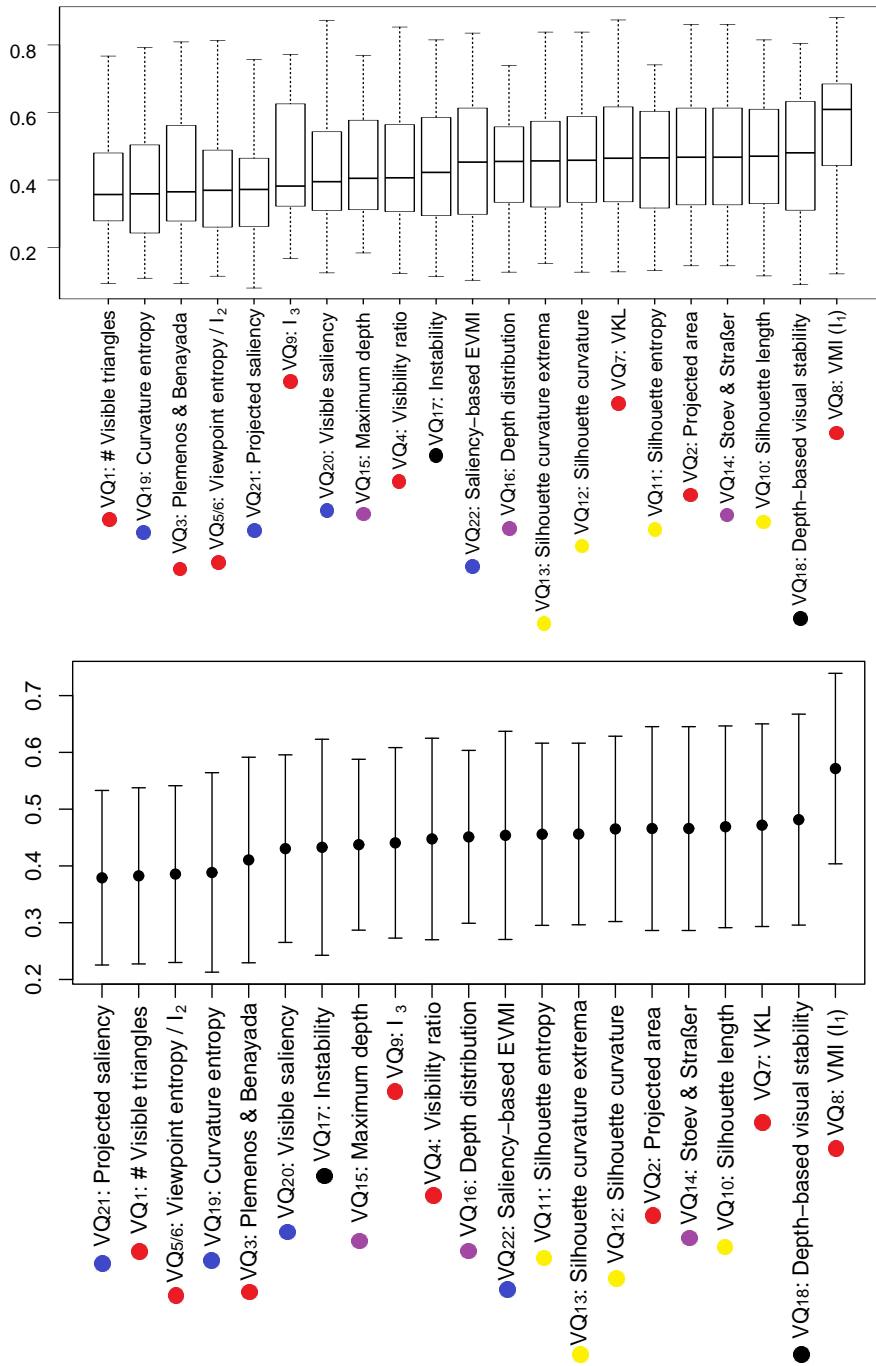


**Figure 1.** Set of best views for the armadillo, (a,d), the cow, (b,e), and the dragon, (c,f), selected by the 26 human subjects in the Dutagaci et al. [14] benchmark.

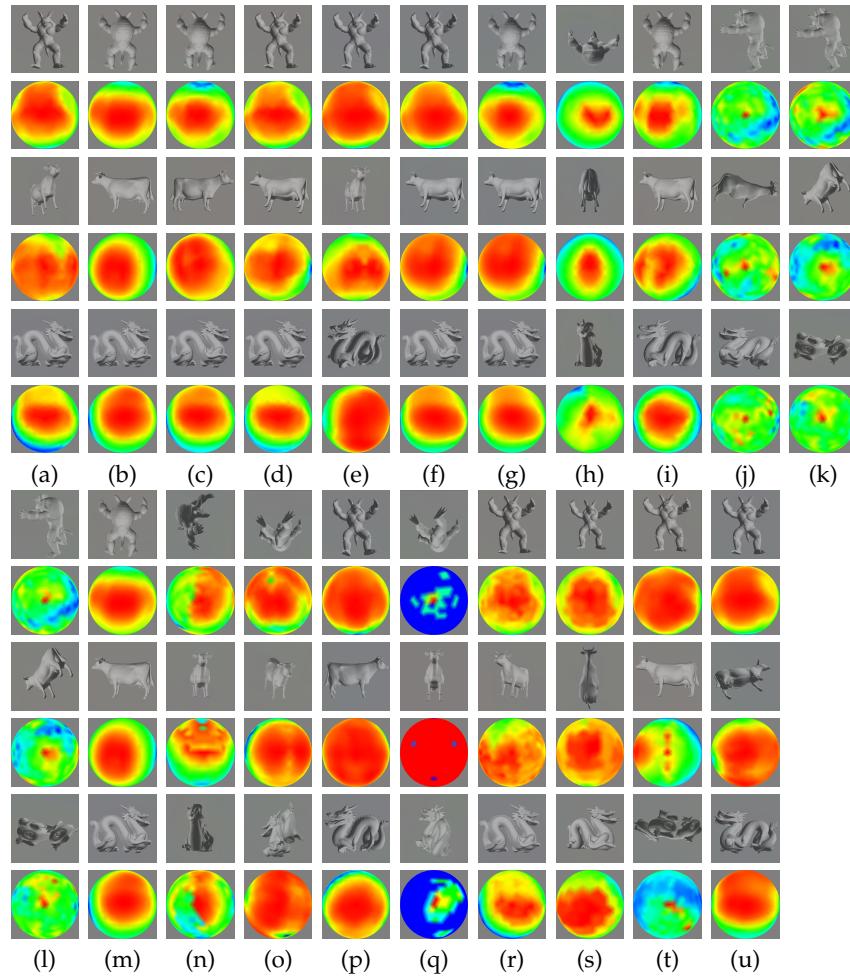
Figure 3 (from column (a) to column (u)) shows the best view and the corresponding viewpoint sphere obtained with the following viewpoint quality measures: (a) number of visible triangles,

(b) projected area, (c) Plemenos and Benayada, (d) visibility ratio, (e) viewpoint entropy/ $I_2$ , (f) viewpoint Kullback–Leibler distance, (g) viewpoint mutual information (or  $I_1$ ), (h)  $I_3$ , (i) silhouette length, (j) silhouette entropy, (k) silhouette curvature, (l) silhouette curvature extrema, (m) Stoev and Straßer, (n) maximum depth, (o) depth distribution, (p) instability, (q) depth-based visual stability, (r) curvature entropy, (s) visual saliency, (t) projected saliency, and (u) saliency-based EVMI. Rows (i), (iii), and (v) show, respectively, the best views of the armadillo, the cow, and the dragon, and rows (ii), (iv), and (vi) show the corresponding viewpoint sphere from the selected viewpoint. The sphere of viewpoints is represented by a color map, where red and blue colors correspond, respectively, to the best and worst viewpoints in terms of the corresponding viewpoint quality measures. From the different distributions, we can see the preferred and unfavored regions, the transition between them, and also the stability of the measure with respect to small viewpoint variations.

We evaluate the set of measures presented in Section 3 with Dutagaci’s benchmark [14]. This benchmark uses the most informative view of 68 models chosen by 26 human subjects. An error between 0 and 1 and the average for all the models can be computed using the benchmark. In Figure 2, we show the box plot ordered by median (top) and the mean +/- the standard deviation ordered by mean (bottom) of the error of the models for each method. We also mark the category of each measure with a color: area attribute (red), silhouette attribute (yellow), depth attribute (purple), stability attribute (black), and surface curvature attribute (blue). Observe that, if we rank the measures in terms of mean and median, the sets of the five best ones are the same: projected saliency [35], the number of visible triangles [19], viewpoint entropy and  $I_2$  [21,41], curvature entropy [13], and Plemenos and Benayada [19]. Observe also that the five best measures belong to two categories: area attributes and surface curvature attributes. In contrast, the measures from the silhouette attributes category perform poorly. One reason for this could be returning to the idea that we represent objects in terms of volumetric primitives [55,56]. In this regard, area attributes and curvature could allow for the reconstruction of (or access to) higher-order representations than 2D image based properties. One could argue that it is not sufficient for the outline of an object to access mental representations of objects, but, rather, the outlines or properties that allow for the identification of element parts.



**Figure 2.** Box plot ordered by median (Top) and mean +/- standard deviation (Bottom) of the error of each method running the Dutagaci et al. [14] benchmark that checks 68 different models. The attribute category is marked with a color dot: area (red), silhouette (yellow), depth (purple), stability (black), and surface curvature (blue).



**Figure 3.** The best view and the corresponding sphere of viewpoints of three models using different methods: (a)  $VQ_1$ , # visible triangles; (b)  $VQ_2$ , projected area; (c)  $VQ_3$ , Plemenos and Benayada; (d)  $VQ_4$ , visibility ratio; (e)  $VQ_5$ , viewpoint entropy /  $VQ_6$ ,  $I_2$ ; (f)  $VQ_7$ , viewpoint Kullback–Leibler; (g)  $VQ_8$ , viewpoint mutual information ( $I_1$ ); (h)  $VQ_9$ ,  $I_3$ ; (i)  $VQ_{10}$ , silhouette length; (j)  $VQ_{11}$ , silhouette entropy; (k)  $VQ_{12}$ , silhouette curvature; (l)  $VQ_{13}$ , silhouette curvature extrema; (m)  $VQ_{14}$ , Stoev and Straßer; (n)  $VQ_{15}$ , maximum depth; (o)  $VQ_{16}$ , depth distribution; (p)  $VQ_{17}$ , instability; (q)  $VQ_{18}$ , depth-based visual stability; (r)  $VQ_{19}$ , curvature entropy; (s)  $VQ_{20}$ , visual saliency; (t)  $VQ_{21}$ , projected saliency; and (u)  $VQ_{22}$ , saliency-based EVMI

## 5. Applications

We present here some applications of the viewpoint quality measures of Section 3 to other fields of research. In Table 3, for each reference, we specify the measure(s) used or the measures the reference is inspired by, and the field of application. The fields of application considered in Table 3 are scene exploration and camera placement (SE/CP), image-based modeling and rendering (IBMR), scientific visualization (SV), shape retrieval (SR), and mesh simplification (MS). We only review the papers related to the measures presented in Section 3. Note also that some of the measures in Table 3 might not fully match the ones introduced in Section 3, but they are as closely related as to be considered under the same token.

Barral et al. [36,57] apply viewpoint quality measures to compute an efficient exploratory path for the visual understanding of a scene. Vázquez and Sbert [58] present a method for the automatic exploration of indoor scenes. They take into account the increase of information in terms of viewpoint entropy to decide the next position and orientation of the camera. Andújar et al. [59] present an algorithm for the automatic exploration of a scene. First, a cell-and-portal detection method

identifies the over-all structure of the scene; second, an entropy-based measurement algorithm is used to identify the cells that are worth visiting, and third, a path is built that traverses all the relevant cells. Feixas et al. [35] present two object exploration algorithms based on viewpoint mutual information. In the first algorithm (guided tour), the path visits a set of  $N$  preselected best views, which ensures a good exploration of the object. In the second algorithm (exploratory tour), the successive viewpoints are selected using the maximum novelty criterion with respect to the parts seen of the object. Ozaki et al. [60] use viewpoint entropy to automatically generate a smooth movement of a camera to follow a subject. Serin et al. [61] present a viewpoint entropy-based approach to navigate over a 3D terrain. Best viewpoints for extracted subregions are calculated with a greedy  $N$ -best view selection algorithm.

Massios and Fisher [62] use the next best view for the reconstruction of a 3D object using a laser range scanner with the minimum number of viewpoints. Fleishman et al. [63] use the projected area to compute a minimum set of viewpoint inside a walking zone for image-based modeling. Vázquez et al. [37] use viewpoint entropy to minimize the number of images used for image-based rendering.

Bordoloi and Shen [48] compute the instability and the viewpoint entropy in volume rendering to select the best and the  $N$  best views of volumetric data. They also apply them to time-varying data. Takahashi et al. [64] apply viewpoint entropy to volume visualization by decomposing an entire volume into a set of feature components. Ji and Shen [65] implement a time-varying view for time-varying volumes in order to maximize the amount of information seen each moment with smooth transitions. Viola et al. [43] use viewpoint mutual information and a similar saliency-based EVMI to select the most expressive view for a specific focus of attention. When the user changes the focus of attention, the viewpoint is changed smoothly. Ruiz et al. [66] use a variation of projected saliency with the voxel information to compute the best viewpoint of a volume data set. Ruiz et al. [67] apply viewpoint Kullback–Leibler to compute automatic transfer functions. Itoh et al. [68] use a variant of viewpoint entropy for automatically selecting optimal viewpoints for visualizing spatio-temporal characteristics of trajectories on a crossroad using a space time cube. Tao et al. [69] apply viewpoint mutual information to 3D flow visualization to select best viewpoints, to decide how to cluster the streamlines and to create a camera path for automatic flow field exploration. Lee et al. [70] use normalized Shannon entropy to create a volumetric scalar entropy field to measure the complexity of a vector field. Maximum intensity projection of this volume is then used to obtain the maximum entropy viewpoints. Vázquez et al. [71] apply the use of minimum and maximum viewpoint entropy to the visualization of molecular structures to study their chemical and physical properties. Sarikaya et al. [72] use viewpoint selection techniques to identify features of interest on protein surfaces and to explore them efficiently.

González et al. [73] use viewpoint mutual information to compute the similarity between two 3D objects. Eitz et al. [74] use best view selection to retrieve a 3D object from a database given a 2D sketch. Li et al. [75] use viewpoint entropy to cluster the viewpoints and retrieve the most similar 3D object from a database given a 2D sketch. Bonaventura et al. [76] use viewpoint mutual information and information  $I_2$  to compute the similarity between two 3D objects.

Castelló et al. [77] use viewpoint entropy, a variation of viewpoint Kullback–Leibler [78], viewpoint mutual information [79] and Tsallis generalization of viewpoint entropy and viewpoint mutual information [80], as measures for mesh simplification.



## 6. Conclusions

In this survey, we have reviewed a set of twenty-two measures for viewpoint selection, where eleven of these measures were not reviewed previously. We have extended a previous existing classification of viewpoint measures by Secord et al. [15], and we have implemented and compared them in a single framework, so as to allow for a fair comparison. As ground truth, we have used the Dutagaci et al. [14] user evaluation database. Our public framework allows for easily including any new measure for comparison, or use another database as ground-truth. The results short-listed five measures that effectively represent the viewpoint preferences of the users, between them three measures closely related to information theory. Finally, we have also presented the application fields that the different measures have been employed in, given that their utility could vary according to the purposes that they were designed for. In the future work, we will analyze the combination of some of the measures presented here. For instance, it is worth investigating a convex linear combination of the information-theoretic measures  $I_1$ ,  $I_2$ ,  $I_3$ , as it also provides a decomposition of mutual information. In addition, application of viewpoint measures to other fields such as augmented reality and 3D eye-tracking can be considered.

**Author Contributions:** X.B. wrote the code for the comparison, and an earlier draft of the paper, while being a PhD student advised by M.F. and M.S., who both participated in the design and corrected the first version of the paper, and collaborated in writing subsequent versions. L.C. and C.W. took care of the discussion and writing of the perception aspects of viewpoint selection, and commented and revised the several versions of the paper.

**Acknowledgments:** This work has been partially funded by grant TIN2016-75866-C3-3-R from the Spanish Government, grant 2017-SGR-1101 from Catalan Government and by the National Natural Science Foundation of China (Nos. 61571439, 61471261 and 61771335). Lewis Chuang was supported by the German Research Foundation (DFG) within project C03 of SFB/Transregio 161, and Christian Wallraven by grant 2017-M3C7A1041817 of the National Research Foundation of Korea.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Peters, G. Theories of Three-Dimensional Object Perception—A Survey. *Recent Res. Dev. Pattern Recognit.* **2000**, *1*, 179–197.
2. Biederman, I. Recognition-by-components: A theory of human image understanding. *Psychol. Rev.* **1987**, *94*, 115–147. [[CrossRef](#)] [[PubMed](#)]
3. Koenderink, J.J.; van Doorn, A.J. The internal representation of solid shape with respect to vision. *Biol. Cybern.* **1979**, *32*, 211–216. [[CrossRef](#)] [[PubMed](#)]
4. Edelman, S.; Bültlhoff, H.H. Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vis. Res.* **1992**, *32*, 2385–2400. [[CrossRef](#)]
5. Bültlhoff, H.H.; Edelman, S.Y.; Tarr, M.J. How are three-dimensional objects represented in the brain? *Cereb. Cortex* **1995**, *5*, 247–260. [[CrossRef](#)] [[PubMed](#)]
6. Tarr, M.J.; Bültlhoff, H.H.; Zabinski, M.; Blanz, V. To what extent do unique parts influence recognition across changes in viewpoint? *Psychol. Sci.* **1997**, *8*, 282–289. [[CrossRef](#)]
7. Logothetis, N.K.; Pauls, J. Psychophysical and Physiological Evidence for Viewer-centered Object Representations in the Primate. *Cereb. Cortex* **1995**, *5*, 270–288. [[CrossRef](#)] [[PubMed](#)]
8. Palmer, S.E.; Rosch, E.; Chase, P. Canonical perspective and the perception of objects. In *Attention and Performance IX*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1981; pp. 135–151.
9. Harman, K.L.; Humphrey, G.K.; Goodale, M.A. Active manual control of object views facilitates visual recognition. *Curr. Biol.* **1999**, *9*, 1315–1318. [[CrossRef](#)]
10. Perrett, D.I.; Harries, M.H. Characteristic views and the visual inspection of simple faceted and smooth objects: 'Tetrahedra and potatoes'. *Perception* **1988**, *17*, 703–720. [[CrossRef](#)] [[PubMed](#)]
11. Perrett, D.I.; Harries, M.H.; Looker, S. Use of preferential inspection to define the viewing sphere and characteristic views of an arbitrary machined tool part. *Perception* **1992**, *21*, 497–515. [[CrossRef](#)] [[PubMed](#)]
12. Blanz, V.; Tarr, M.J.; Bültlhoff, H.H. What object attributes determine canonical views? *Perception* **1999**, *28*, 575–600. [[CrossRef](#)] [[PubMed](#)]

13. Polonsky, O.; Patané, G.; Biasotti, S.; Gotsman, C.; Spagnuolo, M. What's in an image? *Vis. Comput.* **2005**, *21*, 840–847. [[CrossRef](#)]
14. Dutagaci, H.; Cheung, C.P.; Godil, A. A benchmark for best view selection of 3D objects. In Proceedings of the ACM Workshop on 3D Object Retrieval, Firenze, Italy, 25 October 2010; pp. 45–50.
15. Secord, A.; Lu, J.; Finkelstein, A.; Singh, M.; Nealen, A. Perceptual models of viewpoint preference. *ACM ToG* **2011**, *30*, 109:1–109:12. [[CrossRef](#)]
16. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.* **1954**, *61*, 183–193. [[CrossRef](#)] [[PubMed](#)]
17. Connolly, C.I. The determination of next best views. In Proceedings of the IEEE International Conference on Robotics and Automation, St. Louis, MO, USA, 25–28 March 1985; Volume 2, pp. 432–435.
18. Kamada, T.; Kawai, S. A simple method for computing general position in displaying three-dimensional objects. *Comput. Vis. Graph. Image Proc.* **1988**, *41*, 43–56. [[CrossRef](#)]
19. Plemenos, D.; Benayada, M. Intelligent Display Techniques in Scene Modelling. New Techniques to Automatically Compute Good Views. In Proceedings of the International Conference GraphiCon'96, St Petersbourg, Russia, 1 July 1996.
20. Arbel, T.; Ferrie, F.P. Viewpoint selection by navigation through entropy maps. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 1, pp. 248–254.
21. Vázquez, P.P.; Feixas, M.; Sbert, M.; Heidrich, W. Viewpoint Selection Using Viewpoint Entropy. In Proceedings of the Vision Modeling and Visualization Conference, Stuttgart, Germany, 21–23 November 2001; pp. 273–280.
22. Weinshall, D.; Werman, M. On view likelihood and stability. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 97–108. [[CrossRef](#)]
23. Stoev, S.L.; Straßer, W. A case study on automatic camera placement and motion for visualizing historical data. In Proceedings of the IEEE Visualization '02, IEEE Computer Society, Washington, DC, USA, 27 October–1 November 2002; pp. 545–548.
24. Yamauchi, H.; Saleem, W.; Yoshizawa, S.; Karni, Z.; Belyaev, A.G.; Seidel, H.P. Towards Stable and Salient Multi-View Representation of 3D Shapes. In Proceedings of the IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06), Matsushima, Japan, 14–16 June 2006; pp. 265–270.
25. Itti, L.; Koch, C.; Niebur, E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1254–1259. [[CrossRef](#)]
26. Borji, A.; Itti, L. State-of-the-Art in Visual Attention Modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 185–207. [[CrossRef](#)] [[PubMed](#)]
27. Lee, C.H.; Varshney, A.; Jacobs, D.W. Mesh saliency. *ACM ToG* **2005**, *24*, 659–666.
28. Gal, R.; Cohen-Or, D. Salient geometric features for partial shape matching and similarity. *ACM Trans. Graph. (TOG)* **2006**, *25*, 130–150. [[CrossRef](#)]
29. Becker, M.W.; Pashler, H.; Lubin, J. Object-intrinsic oddities draw early saccades. *J. Exp. Psychol. Hum. Percept. Perform.* **2007**, *33*, 20–30. [[CrossRef](#)] [[PubMed](#)]
30. Koulieris, G.A.; Drettakis, G.; Cunningham, D.; Mania, K. C-LOD: Context-aware Material Level-Of-Detail applied to Mobile Graphics. *Comput. Graph. Forum* **2014**, *33*, 41–49. [[CrossRef](#)]
31. Gooch, B.; Reinhard, E.; Moulding, C.; Shirley, P. Artistic Composition for Image Creation. In *Rendering Techniques*; Springer: London, UK, 2001; pp. 83–88.
32. Fu, H.; Cohen-Or, D.; Dror, G.; Sheffer, A. Upright orientation of man-made objects. *TOG* **2008**, *27*, 42:1–42:7. [[CrossRef](#)]
33. Zusne, L. *Visual Perception of Form*; Academic Press: Cambridge, MA, USA, 1970.
34. Podolak, J.; Shilane, P.; Golovinskiy, A.; Rusinkiewicz, S.; Funkhouser, T. A planar-reflective symmetry transform for 3D shapes. *TOG* **2006**, *25*, 549–559. [[CrossRef](#)]
35. Feixas, M.; Sbert, M.; González, F. A unified information-theoretic framework for viewpoint selection and mesh saliency. *ACM Trans. Appl. Percept.* **2009**, *6*, 1–23. [[CrossRef](#)]
36. Barral, P.; Dorme, G.; Plemenos, D. Visual understanding of a scene by automatic movement of a camera. In Proceedings of the International Conference GraphiCon'99, Moscow, Russia, 26 August 1999.
37. Vázquez, P.P.; Feixas, M.; Sbert, M.; Heidrich, W. Automatic View Selection Using Viewpoint Entropy and its Applications to Image-based Modelling. *Comput. Graph. Forum* **2003**, *22*, 689–700. [[CrossRef](#)]

38. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley-Interscience: New York, NY, USA, 1991.
39. Yeung, R.W. *Information Theory and Network Coding*; Springer: Berlin, Germany, 2008.
40. Deweese, M.R.; Meister, M. How to measure the information gained from one symbol. *Netw. Comput. Neural Syst.* **1999**, *10*, 325–340. [CrossRef]
41. Bonaventura, X.; Feixas, M.; Sbert, M. Viewpoint Information. In Proceedings of the 21st GraphiCon International Conference on Computer Graphics and Vision, Moscow, Russia, 26 September 2011; pp. 16–19.
42. Sbert, M.; Plemenos, D.; Feixas, M.; González, F. Viewpoint Quality: Measures and Applications. In Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging, Girona, Spain, 18–20 May 2005; pp. 185–192.
43. Viola, I.; Feixas, M.; Sbert, M.; Gröller, M.E. Importance-Driven Focus of Attention. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12*, 933–940. [CrossRef] [PubMed]
44. Butts, D.A. How much information is associated with a particular stimulus? *Netw. Comput. Neural Syst.* **2003**, *14*, 177–187. [CrossRef]
45. Page, D.L.; Koschan, A.F.; Sukumar, S.R.; Roui-Abidi, B.; Abidi, M.A. Shape Analysis Algorithm Based on Information Theory. In Proceedings of the IEEE International Conference on Image Processing (ICIP’03), Barcelona, Spain, 14–18 September 2003; Volume 1, pp. 229–232.
46. Vieira, T.; Bordignon, A.; Peixoto, A.; Tavares, G.; Lopes, H.; Velho, L.; Lewiner, T. Learning Good Views through Intelligent Galleries. *Comput. Graph. Forum* **2009**, *28*, 717–726. [CrossRef]
47. Burbea, J.; Rao, C.R. On the Convexity of some Divergence Measures Based on Entropy Functions. *IEEE Trans. Inf. Theory* **1982**, *28*, 489–495. [CrossRef]
48. Bordoloi, U.D.; Shen, H.W. View selection for volume rendering. In Proceedings of the IEEE Visualization (VIS 05), Minneapolis, MN, USA, 23–28 Octomber 2005; pp. 487–494.
49. Vázquez, P.P. Automatic view selection through depth-based view stability analysis. *Vis. Comput.* **2009**, *25*, 441–449. [CrossRef]
50. Taubin, G. Estimating the tensor of curvature of a surface from a polyhedral approximation. In Proceedings of the IEEE International Conference on Computer Vision, Washington, DC, USA, 20–23 June 1995; pp. 902–907.
51. Sokolov, D.; Plemenos, D. Viewpoint quality and scene understanding. In Proceedings of the 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST, Pisa, Italy, 8–11 November 2005; pp. 67–73.
52. Bonaventura, X.; Feixas, M.; Sbert, M. Information measures for object understanding. *Signal Image Video Proc.* **2013**, *7*, 467–478. [CrossRef]
53. Serin, E.; Sumengen, S.; Balcisoy, S. Representational image generation for 3D objects. *Vis. Comput.* **2013**, *29*, 675–684. [CrossRef]
54. A Framework for Viewpoint Selection. Available online: <http://github.com/limdor/quoniam> (accessed on 24 March 2018).
55. Hayworth, K.J.; Biederman, I. Neural evidence for intermediate representations in object recognition. *Vis. Res.* **2006**, *46*, 4024–4031. [CrossRef] [PubMed]
56. Hummel, J.E.; Biederman, I. Dynamic binding in a neural network for shape recognition. *Psychol. Rev.* **1992**, *99*, 480–517. [CrossRef] [PubMed]
57. Barral, P.; Dorme, G.; Plemenos, D. Scene understanding techniques using a virtual camera. In *Proceedings of Eurographics 2000, Short Presentations, Rendering and Visibility*; Eurographics Association: Interlaken, Switzerland, 2000.
58. Vázquez, P.P.; Sbert, M. Automatic indoor scene exploration. In Proceedings of the 6th International Conference on Computer Graphics and Artificial Intelligence (3IA), Limoges, France, 14–15 May 2003; pp. 13–24.
59. Andújar, C.; Vázquez, P.P.; Fairén, M. Way-Finder: guided tours through complex walkthrough models. *Comput. Graph. Forum* **2004**, *23*, 499–508. [CrossRef]
60. Ozaki, M.; Gobeawan, L.; Kitaoka, S.; Hamazaki, H.; Kitamura, Y.; Lindeman, R.W. Camera movement for chasing a subject with unknown behavior based on real-time viewpoint goodness evaluation. *Vis. Comput.* **2010**, *26*, 629–638. [CrossRef]
61. Serin, E.; Adali, S.H.; Balcisoy, S. Automatic path generation for terrain navigation. *Comput. Graph.* **2012**, *36*, 1013–1024. [CrossRef]

62. Massios, N.A.; Fisher, R.B. A Best Next View Selection Algorithm Incorporating a Quality Criterion. In *Proceedings of the British Machine Vision Conference*; BMVA Press: Southampton, UK, 1998; pp. 78.1–78.10.
63. Fleishman, S.; Cohen-Or, D.; Lischinski, D. Automatic Camera Placement for Image-Based Modeling. *Comput. Graph. Forum* **2000**, *19*, 101–110. [[CrossRef](#)]
64. Takahashi, S.; Fujishiro, I.; Takeshima, Y.; Nishita, T. A Feature-Driven Approach to Locating Optimal Viewpoints for Volume Visualization. In Proceedings of the IEEE Visualization 2005, Minneapolis, MN, USA, 23–28 October 2005; pp. 495–502.
65. Ji, G.; Shen, H.W. Dynamic View Selection for Time-Varying Volumes. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12*, 1109–1116. [[PubMed](#)]
66. Ruiz, M.; Boada, I.; Feixas, M.; Sbert, M. Viewpoint information channel for illustrative volume rendering. *Comput. Graph.* **2010**, *34*, 351–360. [[CrossRef](#)]
67. Ruiz, M.; Bardera, A.; Boada, I.; Viola, I.; Feixas, M.; Sbert, M. Automatic Transfer Functions Based on Informational Divergence. *IEEE Trans. Vis. Comput. Graph.* **2011**, *17*, 1932–1941. [[CrossRef](#)] [[PubMed](#)]
68. Itoh, M.; Yokoyama, D.; Toyoda, M.; Kitsuregawa, M. Optimal viewpoint finding for 3D visualization of spatio-temporal vehicle trajectories on caution crossroads detected from vehicle recorder big data. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 3426–3434.
69. Tao, J.; Ma, J.; Wang, C.; Shene, C.K. A Unified Approach to Streamline Selection and Viewpoint Selection for 3D Flow Visualization. *IEEE Trans. Vis. Comput. Graph.* **2013**, *19*, 393–406. [[CrossRef](#)] [[PubMed](#)]
70. Lee, T.Y.; Mishchenko, O.; Shen, H.W.; Crawfis, R. View point evaluation and streamline filtering for flow visualization. In Proceedings of the 2011 IEEE Pacific Visualization Symposium, Hong Kong, China, 1–4 March 2011; pp. 83–90.
71. Vázquez, P.P.; Feixas, M.; Sbert, M.; Llobet, A. Realtime automatic selection of good molecular views. *Comput. Graph.* **2006**, *30*, 98–110. [[CrossRef](#)]
72. Sarikaya, A.; Albers, D.; Mitchell, J.C.; Gleicher, M. Visualizing Validation of Protein Surface Classifiers. *Comput. Graph. Forum* **2014**, *33*, 171–180. [[CrossRef](#)] [[PubMed](#)]
73. González, F.; Feixas, M.; Sbert, M. *View-Based Shape Similarity Using Mutual Information Spheres*; EG Short Papers; Eurographics Association: Prague, Czech Republic, 2007; pp. 21–24.
74. Eitz, M.; Richter, R.; Boubekeur, T.; Hildebrand, K.; Alexa, M. Sketch-based shape retrieval. *TOG* **2012**, *31*, 31:1–31:10. [[CrossRef](#)]
75. Li, B.; Lu, Y.; Johan, H. Sketch-Based 3D Model Retrieval by Viewpoint Entropy-Based Adaptive View Clustering. In *Proceedings of the Sixth Eurographics Workshop on 3D Object Retrieval*; Eurographics Association: Aire-la-Ville, Switzerland, 2013; pp. 49–56.
76. Bonaventura, X.; Guo, J.; Meng, W.; Feixas, M.; Zhang, X.; Sbert, M. 3D shape retrieval using viewpoint information-theoretic measures. *Comput. Anim. Virtual Worlds* **2015**, *26*, 147–156. [[CrossRef](#)]
77. Castelló, P.; Sbert, M.; Chover, M.; Feixas, M. Viewpoint entropy-driven simplification. In Proceedings of the 15th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2007), Plzen, Czech Republic, 29 January–1 February 2007; pp. 249–256.
78. Castelló, P.; Sbert, M.; Chover, M.; Feixas, M. Viewpoint-based simplification using f-divergences. *Inf. Sci.* **2008**, *178*, 2375–2388. [[CrossRef](#)]
79. Castelló, P.; Sbert, M.; Chover, M.; Feixas, M. Viewpoint-driven simplification using mutual information. *Comput. Graph.* **2008**, *32*, 451–463. [[CrossRef](#)]
80. Castelló, P.; González, C.; Chover, M.; Sbert, M.; Feixas, M. Tsallis Entropy for Geometry Simplification. *Entropy* **2011**, *13*, 1805–1828. [[CrossRef](#)]

