Viewpoint Information

Xavier Bonaventura, Miquel Feixas, Mateu Sbert*
Graphics and Imaging Laboratory
University of Girona, Spain
{xavierb, feixas, mateu}@ima.udg.edu

Abstract

In this paper, we present a new perspective to quantify the information associated with a viewpoint. The starting point is twofold: a visibility channel between a set of viewpoints and the polygons of an object, and two specific information measures introduced respectively by DeWeese and Meister (1999) and Butts (2003) to evaluate the significance of stimuli and responses in the neural code. In our approach, these information measures are applied to the visibility channel in order to quantify the information associated with each viewpoint and are compared with both viewpoint entropy and viewpoint mutual information. A number of experiments show the behavior of the proposed measures in best view selection.

Keywords: *Viewpoint selection, Mutual information, Specific information.*

1. INTRODUCTION

In computer graphics, several viewpoint quality measures, such as viewpoint entropy and viewpoint mutual information, have been applied in areas such as best view selection for polygonal models [1, 2], scene exploration [3], and volume visualization [4, 5]. Best view selection is also a fundamental task in object recognition. Many works have demonstrated that the recognition process is view-dependent [6, 7, 8]. On the one hand, Tarr et al. [7] found that "visual recognition may be explained by a view-based theory in which viewpoint-specific representations encode both quantitative and qualitative features". On the other hand, Palmer et al. [6] and Blanz et al. [8] have presented different experiments demonstrating that observers prefer views (called canonical views) that avoid occlusions and that are off-axis (such as a three-quarter viewpoint), salient (the most significant characteristics of an object are visible), stable, and with a large number of visible surfaces.

In this paper, we propose two new viewpoint quality measures that are respectively derived from two different decompositions of mutual information proposed by DeWeese and Meister [9] and Butts [10] in the field of neural systems to quantify the information associated with stimuli and responses. First, we set an information channel between a set of viewpoints and the polygons of an object, and, then, we use those information measures to calculate the information associated with a viewpoint. Experimental results show the performance of these information measures to evaluate the quality of a viewpoint. This paper is organized as follows. In Section 2, we present the most basic information-theoretic measures and different decompositions of mutual information that are applied to quantifying the information associated with stimuli and responses. In Section 3, two new viewpoint information measures are presented. In Section 4, experimental results show the behavior of the proposed measures to select the best views. Finally, in Section 5, our conclusions and future work are presented.

*This work has been funded in part by grant number TIN2010-21089-C03-01 of Spanish Government and grant number 2009-SGR-643 of Generalitat de Catalunya (Catalan Government).

2. INFORMATION THEORY TOOLS

In this section, we present the most basic information measures and also three different ways of decomposing the mutual information between two random variables.

2.1 Basic Information Measures

Let X be a random variable with alphabet $\mathscr{X}$ and probability distribution $\{p(x)\}$, where $p(x) = Pr\{X = x\}$ and $x \in \mathscr{X}$. Likewise, let Y be a random variable taking values y in $\mathscr{Y}$. A communication channel $X \to Y$ between two random variables (input X and output Y) is characterized by a *probability transition matrix* (composed of conditional probabilities) which determines the output distribution given the input distribution [11].

The *Shannon entropy* $H(X)$ of a random variable X is defined by

$$H(X) = -\sum_{x\in\mathscr{X}} p(x)\log p(x). \quad (1)$$

Entropy measures the average *uncertainty* of a random variable X. All logarithms are base 2 and entropy is expressed in bits. The convention that $0\log 0 = 0$ is used. The *conditional entropy* $H(Y|X)$ is defined by

$$H(Y|X) = \sum_{x\in\mathscr{X}} p(x)H(Y|x), \quad (2)$$

where $p(y|x) = Pr[Y = y|X = x]$ is the conditional probability and $H(Y|x) = -\sum_{y\in\mathscr{Y}} p(y|x)\log p(y|x)$ expresses the uncertainty of Y given x. $H(Y|X)$ measures the average uncertainty associated with Y if we know the outcome of X, and $H(X) \geq H(X|Y) \geq 0$.

The *mutual information* $I(X;Y)$ between X and Y is defined by

$$I(X;Y) = H(Y) - H(Y|X) = \sum_{x\in\mathscr{X}} p(x) \sum_{y\in\mathscr{Y}} p(y|x)\log\frac{p(y|x)}{p(y)}. \quad (3)$$

Mutual information expresses the *shared information* or dependence between X and Y. That is, mutual information expresses how much the knowledge of Y decreases the uncertainty of X, or vice versa. It can be seen that $I(X;Y) = I(Y;X) \geq 0$. If X and Y are independent, then $I(X;Y) = 0$.

2.2 Decomposition of Mutual Information

Given a communication channel $X \to Y$, mutual information can be decomposed in different ways to obtain the information associated with a state in $\mathscr{X}$ or $\mathscr{Y}$. Next, we present different definitions of information that have been analyzed in the field of neural systems to investigate the significance associated to stimuli and responses [9, 10].

For random variables S and R, representing an ensemble of stimuli $\mathscr{S}$ and a set of responses $\mathscr{R}$, respectively, mutual information (see Equation 3) is given by

$$I(S;R) = H(R) - H(R|S) \quad (4)$$

$$= \sum_{s\in\mathscr{S}} p(s) \sum_{r\in\mathscr{R}} p(r|s)\log\frac{p(r|s)}{p(r)}, \quad (5)$$

where $p(r|s)$ is the conditional probability of value r known value s, and $p(S) = \{p(s)\}$ and $p(R) = \{p(r)\}$ are the marginal probability distributions of the input and output variables of the channel, respectively. Note that capital letters S and R as arguments of $p(.)$ are used to denote probability distributions.

To quantify the information associated to each stimulus or response, $I(S;R)$ can be decomposed as

$$I(S;R) = \sum_{s\in\mathscr{S}} p(s)I(s;R) \quad (6)$$

$$= \sum_{r\in\mathscr{R}} p(r)I(S;r), \quad (7)$$

where $I(s;R)$ and $I(r;S)$ represent, respectively, the information associated to stimulus s and response r. Thus, $I(S;R)$ can be seen as a weighted average over individual contributions from particular stimuli or particular responses. The definition of the contribution $I(s;R)$ or $I(S;r)$ can be performed in multiple ways, but we present here the three most basic definitions denoted by I_1, I_2 [9], and I_3 [10].

Given a stimulus s, three information measures that fulfill (6) are:

- The *surprise* I_1 can be directly derived from (5), taking the contribution of a single stimulus to $I(S;R)$:

  $$I_1(s;R) = \sum_{r\in\mathscr{R}} p(r|s)\log\frac{p(r|s)}{p(r)}. \quad (8)$$

  This measure expresses the surprise about R from observing s. It can be shown that I_1 is the only positive decomposition of $I(S;R)$ [9]. This positivity can be proven by the fact that $I_1(s;R)$ is the Kullback-Leibler distance [11] between $p(R|s)$ and $p(R)$.

- The *specific information* I_2 [9] can be derived from (4), taking the contribution of a single stimulus to $I(S;R)$:

  $$I_2(s;R) = H(R) - H(R|s) \quad (9)$$
  $$= -\sum_{r\in\mathscr{R}} p(r)\log p(r) + \sum_{r\in\mathscr{R}} p(r|s)\log p(r|s).$$

  This measure expresses the change in uncertainty about R when s is observed. Note that I_2 can take negative values. This means that certain observations s do increase our uncertainty about the state of the variable R.

- The *stimulus-specific information* I_3 (see [10] for a proof):

  $$I_3(s;R) = \sum_{r\in\mathscr{R}} p(r|s)I_2(S;r). \quad (10)$$

  A large value of $I_3(s;R)$ means that the states of R associated with s are very informative in the sense of $I_2(S;r)$. That is, the most informative input values s are those that are related to the most informative output values r.

Similar to the above definitions for a stimulus s, the information associated to a response r could be defined. In the next sections, these information measures will be studied with more detail in the context of a communication channel between viewpoints and polygons.

The properties of positivity and additivity of these measures have been studied in [9, 10]. A measure is additive when the information obtained about X from two observations, $y \in \mathscr{Y}$ and $z \in \mathscr{Z}$, is equal to that obtained from y plus that obtained from z when y is known. While I_1 is always positive and non-additive, I_2 can take negative values but is additive, and I_3 can take negative values and is non additive. Because of the additivity property, DeWeese and Meister [9] prefer I_2 against I_1 since they consider that additivity is a fundamental property of any information measure.

3. VIEWPOINT QUALITY MEASURES

In this section, we present the main elements of the communication channel between viewpoints and polygons, and then we define the viewpoint information measures derived from the measures presented in Section 2.2.

3.1 Visibility Channel

In this section, we review the elements of an information channel between a set of viewpoints and the set of polygons of an object.

In [2], 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 $\mathscr{V}$ and the set of polygons $\mathscr{Z}$ 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. Viewpoints are indexed by v and polygons by z. The three basic elements of the visibility channel are:

- Conditional probability matrix $p(Z|V)$, where each element $p(z|v) = \frac{a_z(v)}{a_t}$ is defined by the normalized projected area of polygon z over the sphere of directions centered at viewpoint v, $a_z(v)$ is the projected area of polygon z at viewpoint v, and a_t is the total projected area of all polygons over the sphere of directions. Conditional probabilities fulfil $\sum_{z\in\mathscr{Z}} p(z|v) = 1$. The background is not taken into account.

- Input distribution $p(V)$, which represents the probability of selecting each viewpoint, is obtained from the normalization of the projected area of the object at each viewpoint. The input distribution can be interpreted as the importance assigned to each viewpoint v.

- Output distribution $p(Z)$, given by $p(z) = \sum_{v\in\mathscr{V}} p(v)p(z|v)$, which represents the average projected area of polygon z.

From this visibility channel, different measures of viewpoint quality, such as viewpoint entropy [1] and viewpoint mutual information [2], have been defined in the past.

3.2 Viewpoint Information Measures

In this section, the information measures I_1, I_2 and I_3 presented in Section 2.2. are applied to the above visibility channel. Although this perspective of analyzing the viewpoint quality is new, it is important to note that I_1 is equivalent to viewpoint mutual information [2] and I_2 has a close relationship with viewpoint entropy [1].

Given the visibility channel $V \rightarrow Z$, the *viewpoint information* is defined in the following three alternative ways:

- From (8), the *viewpoint information* I_1 of a viewpoint v is defined as

  $$I_1(v;Z) = \sum_{z\in\mathscr{Z}} p(z|v)\log\frac{p(z|v)}{p(z)}. \quad (11)$$

  Observe that I_1 coincides with the viewpoint mutual information defined in [2]. The lowest value of I_1 (i.e., $I_1(v;Z) = 0$) would be obtained when $p(Z|v) = p(Z)$. This means that the distribution of projected areas at a given viewpoint ($p(Z|v)$) would coincide with the average distribution of projected areas from all viewpoints ($p(Z)$). In this case, the view is considered maximally *representative*. Thus, while the most surprising views correspond to the highest I_1 values, the most representative ones correspond to the lowest I_1 values. The best viewpoint is defined as the one that has the lowest value of I_1 (i.e., maximum representativeness).

- From (9), the *viewpoint information* I_2 of a viewpoint v is defined as

$$\begin{aligned} I_2(v;Z) &= H(Z) - H(Z|v) \quad (12) \\ &= -\sum_{z\in\mathscr{Z}} p(z)\log p(z) + \sum_{z\in\mathscr{Z}} p(z|v)\log p(z|v). \end{aligned}$$

While the highest value of I_2 would correspond to a viewpoint that could only see one polygon, the lowest value of I_2 would be obtained if a viewpoint could see all polygons with the same projected area. In this case, the view is maximally *diverse*. The best viewpoint is defined as the one that has the lowest value of I_2 (i.e., maximum diversity).

Specific information $I_2(v;Z)$ is closely related to viewpoint entropy, defined as $H(Z|v)$ [1, 2], 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 will essentially 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 (see [2]), while I_2 presents a more stable behavior due to the normalizing effect of $H(Z)$ in (12). The advantage of I_2 against viewpoint entropy could be appreciated in areas such as object recognition and mesh simplification. In the first case, the stable behavior of I_2 would enable us to compare the obtained values for objects with different mesh resolutions and, in the second case, I_2 would take into account the variation of $H(Z)$ in the simplification process.

- From (10), the *viewpoint information* I_3 of a viewpoint v is defined as

$$I_3(v;Z) = \sum_{z\in\mathscr{Z}} p(z|v) I_2(V;z), \quad (13)$$

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

$$\begin{aligned} I_2(V;z) &= H(V) - H(V|z) \quad (14) \\ &= -\sum_{v\in\mathscr{V}} p(v)\log p(v) + \sum_{v\in\mathscr{V}} p(v|z)\log p(v|z). \end{aligned}$$

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.

As we have seen above, $I_1(x;Y)$, $I_2(x;Y)$, and $I_3(x;Y)$ represent three different ways of quantifying the information associated with a viewpoint v. Observe that we consider that the best views correspond to the lowest values of I_1 and I_2, and the highest values of I_3; and the contrary for the worst views. That is, the goodness of a viewpoint is associated with its representativeness (minimum I_1), diversity (minimum I_2), and informativeness (maximum I_3). The word 'informativeness' is used here to express the capability of I_3 to capture information from the polygons of the object. Another aspect to take into account is that the concept of 'best' or 'worst' is relative to the objective we pursue. Thus, for instance, the 'worst' view in the sense of I_2 could be used to select the view with the lowest diversity, such as the one that better shows the structure of a molecule (see [12]).

	Number of polygons	Computational cost
Coffee cup	10732	3526 ms
Horse	43571	3650 ms
Ship	48811	3822 ms
Lady of Elche	51978	3946 ms

Table 1: Number of polygons of the models used and computational cost of the preprocess step for each model in miliseconds.

4. RESULTS

In this section, the behavior of I_1, I_2, and I_3 is analyzed. To calculate these measures, we need to obtain the projected area of every poligon for every viewpoint, and these areas will enable us to obtain the probabilities of the visibility channel ($p(V)$, $p(Z|V)$, and $p(Z)$). In this paper, all measures have been computed without taking into account the background, and using a projection resolution of 640×480. In our experiments, all the objects are centered in a sphere of 642 viewpoints built from the recursive discretisation of an icosahedron and the camera is looking at the center of this sphere. To obtain the viewpoint sphere, the smallest bounding sphere of the model is obtained and, then, the viewpoint sphere adopts the same center as the bounding sphere and a radius three times the radius of the bounding sphere.

In Table 1 we show the number of polygons of the models used in this section and the cost of the preprocess step, i.e., the cost of computing the projected areas $a_z(v)$ and a_t. To show the behavior of the measures, the sphere of viewpoints is represented by a color map, where red and blue colors correspond respectively to the best and worst views. Remember that a good viewpoint corresponds to a low value of I_1 and I_2, and to high value of I_3. Our tests were run on a Intel© Core™ i5 430M 2.27GHz machine with 4 GB RAM and an ATI Mobility Radeon™ HD 5470 with 512 MB.

To evaluate the performance of the viewpoint quality measures, four models have been used: a coffee cup, a horse, the Lady of Elche, and a ship. Figure 1 has been organized as follows. Rows (i), (iii) and (v) show, respectively, the best (columns (a-d)) and worst (columns (e-h)) views corresponding to I_1, I_2 and I_3, and rows (ii), (iv), and (vi) show the viewpoint spheres corresponding to the views shown in rows (i), (iii) and (v), respectively.

While the best views selected by I_1 show a global view of the object, the best views obtained by I_2 capture the maximum number of polygons in a balanced way (i.e., with a similar projected area). This means that I_2 has a high dependence of the resolution of the mesh, trying to see the areas with a finer discretization. On the contrary, it has been shown in [2] that I_1 is very robust with respect to the variation of the mesh resolution. The behavior of I_3 is very different of the one of I_1 and I_2 because the view with maximum I_3 tries to see the most informative polygons, that in general are placed in the most occluded, salient, and complex areas of the object. To better appreciate the behavior of I_3, the best and worst views (see row(v)) show the degree of informativeness of each polygon using a thermal scale, from blue (minimum information) to red (maximum information). Thus, it can be easily seen how I_3 selects the views with the highest informativeness. It is also important to note that a similar view can be considered as the best for one measure and the worst for another. See for instance the best and worst view of the coffee cup for I_2 and I_1, respectively (Figures 1(iii.b) and 1(i.f)), and the best and worst view of the horse for I_3 and I_1, respectively (Figures 1(v.c) and 1(i.g)).

5. CONCLUSIONS

In this paper, we have presented a new perspective based on the decomposition of mutual information to study the quality of a viewpoint. Two measures of specific information introduced in the field

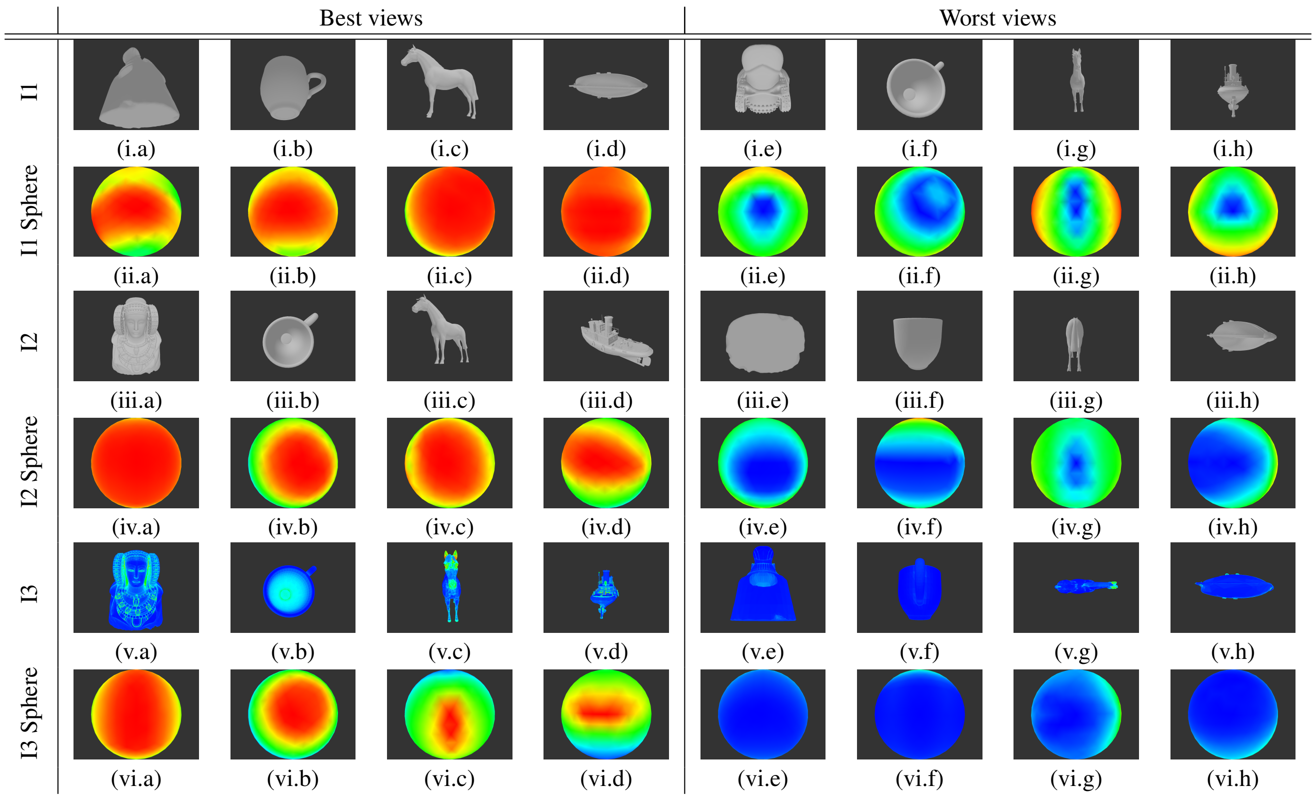

Figure 1: Rows (i), (iii), and (v) show, respectively, the best (a-d) and the worst (e-h) views of four models, obtained with I_1, I_2, and I_3. Rows (ii), (iv), and (vi) show, respectively, the viewpoint spheres corresponding to the views shown in rows (i), (iii), and (v).

of neural systems have been adapted to quantify the information associated with a viewpoint. These measures have been compared with viewpoint entropy and viewpoint mutual information, and different experiments have shown their performance in best view selection. The concepts of surprise, diversity, and informativeness associated with a viewpoint have been also discussed. Further research will be done to analyze the use of the new measures to select *N* best views, to explore a scene, and to compute the information associated with the polygons of an object.

6. REFERENCES

[1] Pere P. Vázquez, Miquel Feixas, Mateu Sbert, and Wolfgang Heidrich, "Viewpoint selection using viewpoint entropy," in *Proceedings of Vision, Modeling, and Visualization 2001*, 2001, pp. 273–280.

[2] Miquel Feixas, Mateu Sbert, and Francisco González, "A unified information-theoretic framework for viewpoint selection and mesh saliency," *ACM Transactions on Applied Perception*, vol. 6, no. 1, pp. 1–23, 2009.

[3] Dmitry Sokolov, Dimitri Plemenos, and Karim Tamine, "Methods and data structures for virtual world exploration.," *The Visual Computer*, vol. 22, no. 7, pp. 506–516, 2006.

[4] Udeepta D. Bordoloi and Han-Wei Shen, "Viewpoint evaluation for volume rendering," in *IEEE Visualization 2005*, 2005, pp. 487–494.

[5] Ivan Viola, Miquel Feixas, Mateu Sbert, and M. Eduard Gröller, "Importance-driven focus of attention," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 933–940, 2006.

[6] S.E. Palmer, E. Rosch, and P. Chase, "Canonical perspective and the perception of objects.," *Attention and Performance IX*, pp. 135–151, 1981.

[7] M.J. Tarr, H.H. Bülthoff, M. Zabinski, and V. Blanz, "To what extent do unique parts influence recognition across changes in viewpoint?," *Psychological Science*, vol. 8, no. 4, pp. 282–289, 1997.

[8] V. Blanz, M.J. Tarr, and H.H. Bülthoff, "What object attributes determine canonical views?," *Perception*, vol. 28, pp. 575–599, 1999.

[9] Michael R. Deweese and Markus Meister, "How to measure the information gained from one symbol," *Network: Computation in Neural Systems*, vol. 10, no. 4, pp. 325–340, November 1999.

[10] Daniel A Butts, "How much information is associated with a particular stimulus?," *Network: Computation in Neural Systems*, vol. 14, pp. 177–187, 2003.

[11] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.

[12] Pere P. Vázquez, Miquel Feixas, Mateu Sbert, and Antoni Llobet, "Realtime automatic selection of good molecular views," *Computers & Graphics*, vol. 30, no. 1, pp. 98–110, 2006.