Institut für Computergraphik und Algorithmen

Technische Universität Wien

Karlsplatz 13/186/2
A-1040 Wien
AUSTRIA

Tel: +43 (1) 58801-18601
Fax: +43 (1) 58801-18698

Institute of Computer Graphics and Algorithms

Vienna University of Technology

email:
technical-report@cg.tuwien.ac.at

other services:
http://www.cg.tuwien.ac.at/
ftp://ftp.cg.tuwien.ac.at/

TECHNICAL REPORT

Importance-Driven Focus of Attention

Ivan Viola   Miquel Feixas   Mateu Sbert†   Meister Eduard Gröller*

TR-186-2-06-02
April 2006

Keywords: illustrative visualization, volume visualization, interacting with volumetric datasets, optimal viewpoint estimation, focus+context techniques

Importance-Driven Focus of Attention

Ivan Viola* Miquel Feixas† Mateu Sbert† Meister Eduard Gröller*

*Institute of Computer Graphics and Algorithms, Vienna University of Technology, Austria
†Institute of Informatics and Applications, University of Girona, Spain

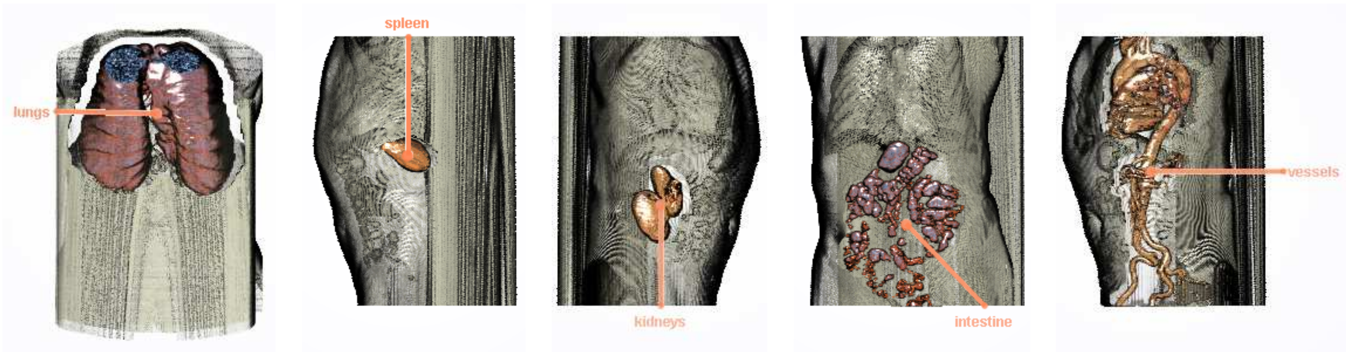

Figure 1: Focus of attention applied to visual inspection of organs within the human torso.

Abstract

This paper introduces a concept for automatic focusing on features within a volumetric data set. The user selects a focus, i.e., object of interest, from a set of pre-defined features. Our system automatically determines the most expressive view on this feature. An optimal viewpoint is estimated by a novel information-theoretic framework which is based on mutual information measure. Viewpoints change smoothly by switching the focus from one feature to another one. This mechanism is controlled by changes in the importance distribution among features in the volume. The highest importance is assigned to the feature in focus. Apart from viewpoint selection, the focusing mechanism also steers visual emphasis by assigning a visually more prominent representation. To allow a clear view on features that are normally occluded by other parts of the volume, the focusing also incorporates cut-away views.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms; I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms;

Keywords: illustrative visualization, volume visualization, interacting with volumetric datasets, optimal viewpoint estimation, focus+context techniques

*{viola | meister}@cg.tuwien.ac.at
†{feixas | mateu}@ima.udg.es

1 Introduction

Visualization is an application-oriented research area combining knowledge from various fields of science or daily life and representing it by means of graphical elements. There can be many reasons for visualization [15]: visual analysis and visual presentation of underlying data are the most important ones. The firstly mentioned reason for visualization helps scientists to find relations and correspondence in various natural phenomena. The visual stimulus is the strongest perception cue and visual analysis helps people to *think visually*. The second goal serves as a communication medium and can be motivated by several reasons such as education, infographics, or commercial purposes. The approach of this paper helps in improving visual presentation rather than visual analysis.

Recent developments of 3D scanning modalities such as computed tomography (CT) allows to unveil insights of different species, material, or bodies. One of the most important application areas is medical diagnostic imaging. Besides medical visualization there are other science directions exploiting 3D scanning technology. For example high-resolution CT scanning has been applied to provide Visible Man datasets or recently the mummy data set of Egypt's boy pharaoh Tut. Industrial scanners are used for material quality validation, but can be used for scanning of small species, e.g., insects as well [24]. An interesting example of exploiting 3D scanning technology is the *Digital Morphology* library [5], a large dynamic archive of information on digital morphology of biological specimens.

As medical imaging is an important application area, medical workstations in general feature the broadest spectrum of functionality for handling volumetric data sets. This includes visualization, image processing, measurements, or (semi-)automatic diagnosis estimation. Medical workstations, however, are designed mostly for visual analysis in diagnostic scenarios, rather than for presentation purposes. Medical workstations currently do not include much functionality that is necessary for presentation purposes.

Other areas of science on the other hand are in general not targeted specifically to visual analysis and visual presentation is often more required. One of these examples is the collection of different specimens in the Digital Morphology library. The aspect of visual presentation is also becoming important in communication among medical experts from different domains or between the medical staff and the patient. Therefore functionality for presentation purposes, also strongly related to visual storytelling, will become more important for medical workstations as well. Functionality that is handling volumetric data sets for presentation purposes has been recently discussed in a tutorial on *illustrative visualization* [21]. Some systems incorporating illustrative presentation techniques are shortly discussed in the related work (Section 2).

Current visualization systems for handling volumetric data sets require a lot of expertise from the user. For example many widgets to design a suitable transfer function (mapping tissue density to color and opacity values) are rather unintuitive for the unexperienced user. Our work is motivated by the fact that currently none of the commercially or publicly available visualization systems allows the user high-level interactions such as "Show me this interesting part of the volumetric data set and then show me the next interesting part." Our framework allows an automatic focus of attention on interesting objects. The user's only required (but not limited to) interaction is to select an object of interest from a set of pre-segmented objects. The framework smoothly navigates the view to optimally see the characteristics of the focus object. Additionally, the focus object is visually emphasized for easy discrimination from the context. Example images that illustrate focus of attention for insights of a human torso and human hand dataset are shown in Figures 1 and 2.

One contribution of this paper is the introduction of an information theoretic framework for optimal viewpoint estimation in volumetric data sets with pre-segmented objects. This framework easily integrates the importance of objects within the volumetric data set. Another contribution is a concept of focus of attention for interactive volume visualization. Here an expressive viewpoint is selected in combination with a visually pleasing discrimination of focus from context information. By changing the object of interest, both viewpoint settings and visual parameters are smoothly changing to put emphasis on the newly selected object of interest.

The paper is organized as follows: Section 2 describes previous work related to importance-driven focus of attention. The following Section 3 describes the concept of focusing. Technical details of optimal viewpoint estimation are discussed in Section 4. Interaction aspects of focusing are presented in Section 5. Implementation issues and performance are discussed in Section 6. Finally we draw conclusions and summarize the paper in Section 7.

2 Related Work

Focus of attention has been often used in visualization to catch the user's attention. It has many different occurrences. We will first review relevant previous work in the area of focus+context visualization. The second part of this section reviews recent work on optimal viewpoint estimation as good viewpoint selection is crucial for an effective focus of attention.

The depth of field effect is a focus of attention technique from photography that inspired Kosara et al. [8] to propose a semantic depth of field (SDOF). In their work they have shown that the degree of sharpness determines the speed of drawing human attention in otherwise blurry environments. They have applied their technique in various fields of information visualization. Later on, the authors have designed a user study for a quantitative evaluation of sematic depth of field efficiency [9]. They show that the semantic depth of field is an effective way to draw attention to specific parts. SDOF, however, should be used in combination with other visualization techniques. Additionally the users sometimes felt uncomfortable observing unnaturally blurred parts (e.g., blurred text).

Another focus+context method for displaying volumetric data has been proposed in previous work on importance-driven volume visualization [22, 23]. The importance classification has been introduced for specifying view-dependent visual representations to reveal occluded structures. This is in spirit of cut-away views and ghosted views known from traditional illustrations.

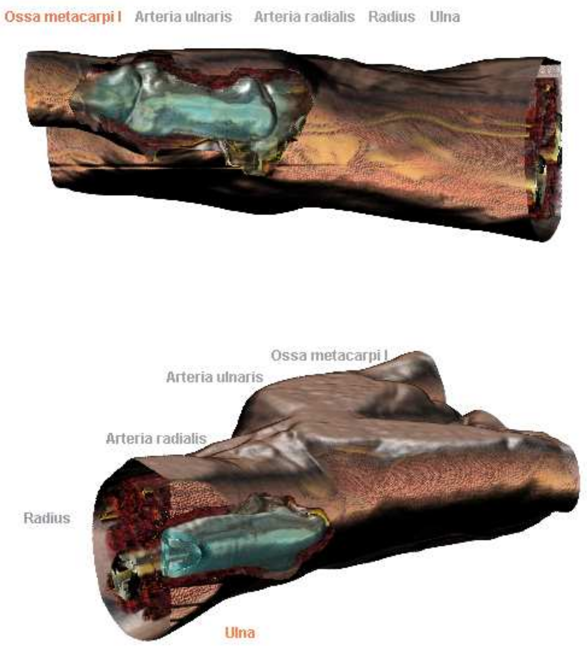

Figure 2: Further examples of visual inspection of organs within the human torso.

Several focus+context techniques have been included into VolumeShop, a publicly available volume visualization system from our group [4]. Besides applying cut-aways and ghosted views, the user can manipulate features of interest in several ways, e.g., displace them from their original spatial location to a part of image space, where otherwise no data is shown. Additionally to focus+context techniques, the user can enrich the visualization by adding textual information to objects which appears as automatically placed labels. The functionality of VolumeShop is intended to provide a tool for presenting and communicating the data being visualized.

Another publicly available visualization system featuring functionality for visual presentation and communication has been proposed by Svahkine et al. [13]. An interesting consideration incorporated in their system is the level of expertise of the user. This has two implications for the design of the system. First, the user interface and widgets are customized according to the user-expertise level. A non-expert user has a very simple user interface allowing limited flexibility, whereas an expert has much higher flexibility with advanced tools such as a transfer function editor. Second, the level of user expertise implies also different visualization results. An easy to understand visualization is targeted to a non-expert user and more *direct* visualization is targeted to the expert.

Viewpoint selection has been applied to several domains in computer graphics, such as scene understanding and virtual exploration [1], molecular visualization [20], image-based modeling [19], volume visualization [3, 14], and mesh saliency [10]. Different measures for viewpoint evaluation have been used in these fields.

Vázquez et al. [18] have defined the *viewpoint entropy* (Equation 13), as a measure for viewpoint quality evaluation. This measure has been designed primarily for polygonal data, where the best viewpoint is defined as the one that has maximum entropy. Taking into account the background information, this technique may be used for indoor and outdoor scenes as well.

Viewpoint entropy for polygonal data has been recently extended to volumetric scalar data [3], by substituting the area visibility distribution by the voxel visibility distribution divided by the voxel importance (*noteworthiness* factor). This work has also suggested information-theoretic measures for clustering views according to similarity using the Jensen-Shannon divergence from information theory (Equation 12). They also suggested an optimal viewpoint estimation scheme for time-varying data.

It has been shown recently by Sbert et al. [12] that viewpoint entropy is very sensitive to triangulation. The maximum entropy is achieved in areas of very fine triangulation. Therefore they propose a new viewpoint-quality measure for polygonal data based on the *Kullback-Leibler distance* (KL) (Equation 11) denoted as viewpoint KL distance (Equation 14). The viewpoint KL distance is interpreted as the distance between the normalized distribution of projected areas and the ideal projection, given by the normalized distribution of the actual areas. In this case, the background is not taken into account. The minimum value 0 is obtained when the normalized distribution of the projected areas is equal to the normalized distribution of the actual areas. Thus, views of high quality correspond to views with minimal KL distance. One drawback of this measure is that many non-visible or poorly visible polygons in a model can distort the quality of the measure.

3 Focusing Considerations

Before going into technical details of our work we would like to *focus* the reader's *attention* on several considerations we have made during designing our framework. To get a clear high-level overview on the framework functionality, we briefly present the processing pipeline. Technical details will follow in Sections 4 and 5.

Focus discrimination: Focus of attention is a visual discrimination of interesting objects from other elements in an image. It is realized through visual emphasis of the object of interest while other objects presented as context are suppressed. In general a discrimination of the focus from the context can be achieved by different *levels of sparseness* in their visual representation [22]. The focus is represented very *densely* while the context gets a more *sparse* visual representation. Levels of sparseness can be designed in many ways. In photography for example, a very effective technique for object discrimination is the sharpness of the object of interest. Very sharp objects are automatically perceived as being in focus, more blurry objects are contextual information. Levels of sparseness are in this case different sharpness levels. Recent studies [9] have shown that for visualization tasks the modulation of sharpness is not very much preferred by users. In volume visualization tasks the depth-of-field from photography may additionally conflict with visual artifacts such as partial volume effects. In this case opacity, color brightness, and saturation can be used to discriminate the most interesting objects from the rest in a much clearer way.

Characteristic view: In addition to visual discrimination, objects in focus have to be shown from a characteristic view where most of the focus structures are perceivable. The most interesting object must not be occluded by less relevant parts. If possible the focus should be in front of other features. In case that the feature of interest is always occluded by other features, cut-away views or other concepts from illustration can be included into the visualization. In this case it is important that the cut-away region does not entirely remove other interesting objects. If possible, only the least relevant objects are cut away. Furthermore a proper orientation of the up-vector of the viewpoint and a proper positioning of the focus to fulfill aesthetical criteria of composition (e.g., rule of thirds [6]) are important to consider in the viewpoint specification. All mentioned aspects indicate that a proper viewpoint setting is important for the focus of attention.

Focusing Pipeline: Some previous work [22, 23] used an explicit importance classification for focus+context visualization inspired by techniques known from traditional illustration. In the following we give an overview on the pipeline of importance-driven focus of attention (for the part on optimal viewpoint estimation see also Figure 3). The pipeline is presented for the visualization of volumetric datasets. The concept is, however, universal and can be applied to various visualization tasks irrespective from the type of the underlying data. The definition of levels of sparseness for visual representations is highly application dependent. They can be defined explicitely by the user, estimated (semi-)automatically [7], or selected from design galleries [11]. A discussion on levels of sparseness is outside the scope of this paper. We concentrate on the estimation of proper viewpoints and on aspects of focusing during user interaction.

Finding a viewpoint where the characteristics of a specific feature are clearly visible is crucial for focus of attention. This naturally requires the estimation of visibility of the feature under specific viewing settings. In our case, i.e., for objects within the volumetric data set, this process is rather time-consuming as it requires ray casting of the whole data set from various viewpoints. Computing the visibility of features on-the-fly during interaction will strongly limit interaction possibilities. The visibility of features depends on their visual representation. For applications where a frequent change of visual representations is not relevant, the visibility estimation can be easily treated as a pre-processing step, which is executed once prior to the user interaction.

In our importance-driven optimal-viewpoint estimation framework we compute the visibility of an object as its contribution on the finally rendered image. This computation is based on the opacity contribution of each voxel belonging to the object. Object visibility is then mapped to a *conditional probability* of the object for a given viewpoint. These values are used for computation of good viewpoints for a given object. We use for this a novel information-theoretic framework combined with object importance information as described in detail in Section 4.

With selecting visual representations of tagged objects and by identifying representative viewpoints, the crucial information to perform on-the-fly focus of attention is available. We use focus of attention as a tool for visually-pleasing *browsing* among a number of interesting structures within the data. Browsing can be in general used for visual presentations of the volumetric data. In our importance-driven focusing framework we also consider additional information about the data. For example we include information about the *up*-vector of the volume (e.g., in the case of the human anatomy, the head is on top and the legs are at the bottom), in order to preserve natural orientations of viewpoints. The object in focus is located in the center of the viewpoint in order to draw the maximal attention of the user. We blend-in textual information as labels to increase the semantic information content. In general, the more information is available the larger the spectrum of possibilities how to realize a pleasing focus of attention. Browsing is realized as a continuous change of the focus of attention. Visual representations and viewpoint settings continuously change to visually emphasize the newly selected object of interest. These changes are driven by changes in the importance distribution among objects. A detailed discussion on browsing through the structures is given in Section 5.

4 Characteristic Viewpoint

In this section, we describe our approach for selecting a characteristic viewpoint for a particular object. First, we determine the visibility of structures within the volumetric data considering their visual representations. Then we use the visibility as input to the new information-theoretic framework. This framework integrates per-object importance classification, which allows to estimate optimal viewpoints for an object within the volume.

4.1 Visibility Estimation

The first step for a viewpoint evaluation is the estimation of per-object visibility. We use a simple scheme for visibility evaluation, taking into account opacity contribution of voxels on the rendered image. The evaluation of the visibility is done in a ray-casting step. For each sample i along a ray r we evaluate its visibility $v(r,i) = v(r,i-1)\alpha(r,i)$, where $\alpha(r,i)$ is the resampled opacity value at the given sample position i. We implement nearest neighbor and linear interpolation resampling schemes. The visibility of a voxel is given as the sum of visibilities of all resampled points the voxel is contributing to in the resampling step. In case of nearest neighbor interpolation we simply sum the ray sample visibilities *belonging* to this voxel. In case of linear interpolation, we perform a linear *distribution* of the ray sample visibility among all eight surrounding voxels.

Each voxel belongs either to a particular feature or it belongs to the *background* volume. The sum of voxel contributions belonging to a particular feature, estimates the visibility of the feature. We are using non-binary object classifications and a particular voxel may contribute to a number of different features simultaneously. The voxel visibility is simply multiplied by a factor that defines how much the voxel contributes to a particular object.

In our focus of attention framework, we also change the visual representation of the object of interest. This means that the visual representation is not constant during the time of interaction. This has to be taken into account while computing visibilities. Therefore we compute the visibility for each *active* object, i.e., object in focus. This means, for each viewpoint we get $(n+1)$ different visibility values for n objects. Each object is set once as active object and once the visibility is computed with no selected active object. When we search for the optimal viewpoint of a particular object, we use those visibilities where this object has been the active object. In this case the object has a different visual representation from the rest of the volume.

One problem that arises when computing the visibility of objects, is that some features may be completely occluded by other features. This is caused by very dense visual settings. This will mean that there is no optimal viewpoint from which the feature is clearly visible, or all viewpoints are equally good or bad. In order to deal with this problem, we have optionally included cutaway views in the visibility estimation. Here the active object is visible from all viewpoints as the volume region in front of this object is not visible at all.

The above described visibility evaluation does not consider the location of features in image space. To draw attention to a feature, it is important that it is located close to the center of the image. To bring the feature into the focus, we give more prominence to rays in the center of the image. We weight each ray's contribution to the visibility of objects and background volume by an image space weight. This weight is largest in the center of the image and is decreasing with the distance from the center.

The overall concept of optimal viewpoint estimation driven by an importance distribution is illustrated in Figure 3. The importance distribution and the visibility of each object for the given visual representations are input parameters of the information-theoretic framework. This framework will be described in detail in the next section.

4.2 Information-Theoretic Framework for Viewpoint Estimation

After the visibility of each object under different visual settings and viewpoints has been computed, the optimal viewpoint estimation can be performed. In this section we describe our approach for finding good viewpoints. Our viewpoint selection approach is using the mutual information of the channel defined between a set of viewpoints and the objects of a volumetric data set. This new measure shows a better behavior and robustness than the previous viewpoint entropy [18]. For more information

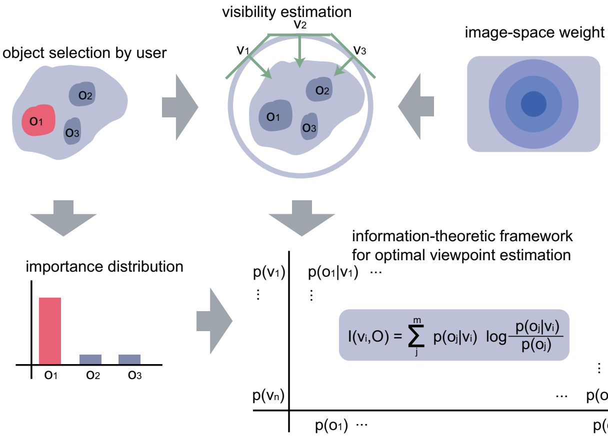

Figure 3: Concept of importance-driven optimal viewpoint estimation.

on information-theoretic measures and previous viewpoint quality measures, please refer to the Appendix section.

Our framework works well for volumetric objects as tagged volume regions. Taking a voxel as a basic object element would lead to very high memory consumption, this will be also the case using previously suggested viewpoint quality measures. The framework naturally integrates per-object importance classification. By changing the importance distribution among objects, the results of viewpoint evaluation also change to have a characteristic view on the feature of highest importance. Setting the importance to be constant for all objects and background volume, characteristic views for the entire volume are achieved.

We formalize the viewpoint selection using a communication channel between two random variables (input and output). The communication channel is characterized by a probability transition matrix which determines the output distribution given the input. After defining a channel, entropy and mutual information can be calculated. The entropy is a measure of the average uncertainty in a random variable and the mutual information is a measure of the dependence between two random variables, i.e., the amount of information one random variable contains about another. While entropy is the self-information of a random variable, mutual information is a special case of a more general quantity called relative entropy, which is a measure of the distance between two probability distributions.

We first define a channel $V \to O$ between a set of viewpoints and the objects of a volumetric data set, represented respectively by the random variables V (input) and O (output). Viewpoints will be indexed by v and objects by o. The marginal probability distribution of V is given by $p(v) = \frac{1}{N_v}$, where N_v is the number of viewpoints, i.e., we assign the same probability to each viewpoint. The conditional (or transition) probabilities $p(o|v)$ are given by the normalized visibility of each object from each viewpoint, i.e., $\sum_{o\in\mathscr{O}} p(o|v) = 1$. Finally, the marginal probability distribution of O is given by

$$p(o) = \sum_{v\in\mathscr{V}} p(v)p(o|v) = \frac{1}{N_v}\sum_{v\in\mathscr{V}} p(o|v), \qquad (1)$$

that expresses the average visibility of each object from the set of viewpoints.

From channel $V \to O$, the *conditional entropy* is given by

$$H(O|V) = -\sum_{v\in\mathscr{V}} p(v) \sum_{o\in\mathscr{O}} p(o|v)\log p(o|v) = \frac{1}{N_v}\sum_{v\in\mathscr{V}} H_v, \qquad (2)$$

where $H_v = -\sum_{o\in\mathscr{O}} p(o|v)\log p(o|v)$ is the *entropy of viewpoint* v (for polygonal data see Equation 13, for volumetric data refer to Bordoloi et al. [3]). Thus, the conditional entropy is the average of all viewpoint entropies.

We now focus our attention on mutual information, that expresses the degree of *dependence* or *correlation* between a set of viewpoints and the data set. The *mutual information* between V and O is given by

$$I(V,O) = \sum_{v\in\mathscr{V}} p(v) \sum_{o\in\mathscr{O}} p(o|v)\log\frac{p(o|v)}{p(o)} = \frac{1}{N_v}\sum_{v\in\mathscr{V}} I(v,O), \qquad (3)$$

where

$$I(v,O) = \sum_{o\in\mathscr{O}} p(o|v)\log\frac{p(o|v)}{p(o)} \qquad (4)$$

is called *viewpoint mutual information* and represents the degree of correlation between the viewpoint v and the set of objects. The quality of a viewpoint is given by the mutual information $I(v,O)$ and the best viewpoint is defined as the one that has minimum mutual information. High values of the measure mean a high dependence between viewpoint v and the set of objects, indicating a highly *coupled* view. On the other hand, low values correspond to a low dependence, allowing for more *representative* views of the data set.

Viewpoint mutual information has the following advantages versus viewpoint entropy. First, the entropy value increases to infinity with the number of voxels and it is highly dependent on the voxel distribution. Thus, an extremely refined mesh attracts the attention of the measure, penalizing big objects in front of small ones. This is not such a problem for volumetric data sets stored on a regular grid, when the basic object element is a voxel. Viewpoint selection evaluation for volumetric data stored on unstructured grid will suffer from this property much more significantly. On the other hand the viewpoint mutual information converges to a finite value when the mesh is infinitely refined and is insensitive to changes in the voxel resolution.

Second, while viewpoint entropy only uses the conditional distribution $p(o|v)$, that is, what is visible from a given point of view, viewpoint mutual information measures how much the distribution $p(o|v)$ differs from the distribution $p(o)$ in the sense of statistical distinctness. Note that $p(o)$ gives us the average visibility of all objects captured from all viewpoints and represents the ideal target, so that the viewpoint mutual information is zero when $p(o|v) = p(o)$. In other words, the viewpoint mutual information considers all the information of the channel. The only objective of viewpoint-entropy maximization is to approach the uniform distribution, without taking into account the degree of visibility of the objects. This behavior of the entropy is independent of weighting the visibility distribution by the importance, as done by Bordoloi and Shen [3].

4.3 Incorporating Importance

We observe that viewpoint mutual information can be rewritten as

$$I(v,O) = KL(p(O|v)|p(O)), \tag{5}$$

where capital letters indicate that $p(O|v)$ is the conditional probability distribution between v and the data set, and $p(O)$ is the marginal probability distribution of O. Thus, $I(v,O)$ can be interpreted as the relative entropy or Kullback-Leibler distance between the visibility distribution of objects from viewpoint v and their average visibility. The less the measure the better the viewpoint, as we approach the ideal target of viewing every object proportional to the average visibility $p(o)$. In this case, $I(v,O)$ would be zero.

Adding importance to our scheme means simply modifying the target function. The ideal viewpoint would be now the one viewing every object proportional to the average visibility times importance. After incorporating importance, the viewpoint mutual information is given by

$$I'(v,O) = \sum_{o \in \mathscr{O}} p(o|v) \log \frac{p(o|v)}{p'(o)}, \tag{6}$$

where

$$p'(o) = \frac{p(o)i(o)}{\sum_{o \in \mathscr{O}} p(o)i(o)} \tag{7}$$

and $i(o)$ is the importance of object o.

4.4 Obtaining Characteristic Viewpoints

Equation 6 defines the viewpoint mutual information with importance classification. This is computed for each viewpoint and for each active object separately (as they have different visual representations, which implies different visibilities). To obtain a set of characteristic views for a given object o, we compute the conditional probabilities of all objects for a given viewpoint. The conditional probability $p(o|v)$ is equal to the normalized visibility, i.e., the visibility of all objects per viewpoint are equal to 1 as described in Section 4.2.

Furthermore we have to compute the marginal probability $p(o)$ from Equation 1. To compute $p'(o)$ we first compute a dot product between the marginal probability vector $(p(o_0), p(o_1), p(o_2), ..., p(o_{m-1}), p(o_m))$ and the importance distribution vector $(i(o_0), i(o_1), i(o_2), ..., i(o_{m-1}), i(o_m))$ where m is the number of objects and o_0 is the background volume. After the sum in the denominator of Equation 7 is computed, all information is available and we can compute the viewpoint mutual information for viewpoint v.

The viewpoint mutual information is computed for every viewpoint and the set of viewpoints with the smallest mutual informations are selected. These computations give us good viewpoints for a particular active object. To compute good viewpoints for another object, we have to take another set of visibilities where the visual emphasis is on the respective object. All values necessary for the viewpoint mutual information can be stored in a set of 2D schemes as shown in Figure 3.

5 Importance-Driven Focusing

The focus of attention requires to display the object of interest from a characteristic view. How to obtain characteristic viewpoints has been described in the previous section. Let's assume we have identified a set of most characteristic viewpoints per object under the given visual representations. Now we will describe in detail how the focus of attention can be used for the visual inspection of tagged volumetric data.

The general idea is to use the importance distribution as a controlling parameter for the focus of attention. We specify a high importance value for the active object (e.g., 100.0). The other objects and the background are assigned a low importance value (e.g., 1.0). By selecting another object to become the active object, the importance of the previously selected active object is continuously decreasing to the value of inactive objects (1.0) and the importance of the newly selected active object is increasing to the maximal value (100.0). The change in the importance distribution is reflected on the viewpoint location and visual representations. This causes a smooth and visually pleasing change of focus of attention to the newly selected active object. We describe these changes in more detail for viewpoints and visual representations separately.

5.1 Viewpoint Transformation

Initially we set the viewpoint to optimally see the entire volume without selecting any active object. All importance values are set to a constant low importance value. For each object we calculate several good characteristic viewpoints, as well as several good viewpoints on the entire volume with no active object selected. All these viewpoints are located on a bounding sphere around the volumetric data set. After selecting an active object, its importance raises and the viewpoint changes to the optimal viewpoint of the active object. As there are several characteristic per-object viewpoints, we have to define which one will be selected. To minimize the viewpoint path, we compute the angle between the current viewpoint's normal vector and the normal vectors of the object's good viewpoints. The viewpoint with the smallest angle is selected. The position of the viewpoint is always on the bounding sphere. The change of position between viewpoints is calculated as a linear transformation from one position to another. Every position of an intermediate viewpoint is then normalized to be located on the bounding sphere surrounding the volume. This has one favorable implication. The viewpoint change starts slowly, has the biggest angle difference in the middle of the viewpoint transformation and slows down before achieving the new optimal viewpoint. This is depicted in Figure 4. The change of viewpoints is parameterized by the importance value of the object in focus. The initial value for this object is low and equal to the importance value of the other objects. After selecting the object as active, the importance increases and defines the position of intermediate viewpoints. When an object's importance value is equal to the maximal value, the viewpoint location is at the desired position.

After achieving the object's characteristic viewpoint, the user can locally change the viewpoint in order to inspect the interesting object from several directions. When this inspection is finished (e.g., mouse button is released), the importance value of the active object is set to a low value again. Increasing the importance brings the viewpoint back to one of the characteristic views.

We use a slightly different concept for changing viewpoints from one active object to another. Instead of a direct viewpoint interpolation from the actual viewpoint to the *closest* new active object's good viewpoint (determined by the angle difference of viewpoint normal vectors), we consider that the viewpoint path visits a viewpoint that gives a general overview on the entire volume. This provides the context information of all structures so the user does not lose its orientation within the volume. From this *context* view, the viewpoint smoothly changes to the optimal viewpoint of the newly selected active object. This way of presenting objects is in the spirit of the navigation on large 2D maps proposed by van Wijk and Nuij [16, 17].

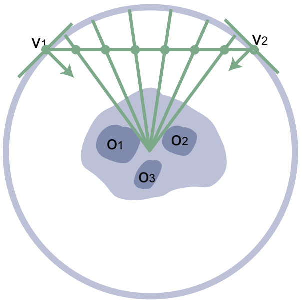

Figure 4: The viewpoint path is calculated as a difference between two viewpoint positions. The path is then normalized onto the bounding sphere, which smooth acceleration and deacceleration in viewpoint change.

In case of switching the viewpoint from one active object to another, the viewpoint change considers three pre-selected viewpoints: the optimal viewpoint of the previous active object, the contextual overview, and the view on the new active object. Instead of the linear transformation discussed above, the position of the viewpoint is changing on a Beziér curve defined by viewpoint positions as three control points [2]. This means that the contextual view on the whole volume is not visited exactly, it is approximated by similar views that also satisfy the goal of providing context. The viewpoint position is again normalized to the unit sphere enclosing the volume. The Beziér curve among three viewpoints is depicted in Figure 5.

In this case we have to select from the two closest optimal viewpoints, i.e., one for the context view and one for the characteristic view on the new active object. We select the viewpoint pair with the smallest sum of angles between the viewpoint normals, which means that the overall path is shortest.

An important consideration in the viewpoint setting with respect to a visually pleasing focusing, is the orientation of the viewpoint *up*-vector. In our implementation we set the viewpoint up-vector to point towards the up-vector of the volume. The up-vector of the volume is

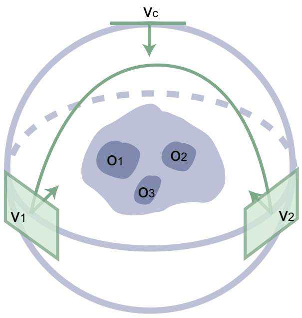

Figure 5: Change between two optimal viewpoint of different objects (v_1 for o_1 and v_2 for o_2). The contextual viewpoint v_c is *nearly* visited and is approximated by the Beziér curve.

defined before the visual inspection for each dataset, together with all other preprocessing steps: defining objects by segmentation, visual representations and good viewpoint estimation. If the viewpoint is located at the *poles*, i.e., the viewpoint normal vector is parallel to the volume top vector, we select another vector to be the viewpoint up-vector. In our implementation we use the volume front vector with inverse orientation so we look at the volume from top-front.

5.2 Visual Representation

A characteristic view is an important part of focus of attention. However without emphasis through visual representation, the focus is still not discriminated from the context objects. Therefore in our framework changes in importance distribution also change the visual appearance of objects. A visual representation basically changes in a similar way as the viewpoint. In this case we do not need to calculate a path. We select the appropriate level of sparseness in the visual representation. In our implementation we define the visual representation of inactive and active objects before the visibility calculation. These visual representations can be linearly interpolated for example. In our focusing pipeline we use a discontinuous change in the visual representation as this abrupt change attracts an observer's attention much stronger. While the viewpoint moves from the previous active selection towards the *context* view (the importance of the previous active object is decreasing), the previous active object is still visually emphasized. After reaching the context viewpoint, the visual representation of the previously active object is suppressed and the new active object is visually emphasized (the importance of the newly selected active object is increasing).

In addition to changes in the visual representation, we incorporate cut-away views. The level of *ghosting* in front of the interesting feature is also driven by importance changes. In this case we do not employ abrupt changes, but the level of ghosting changes smoothly. This means the ghosting level is increasing with decreasing importance of the previously active object and is decreasing with increasing importance of the new active object. When the optimal view is reached, the ghosting level is maximal, i.e., features in front of the active object are completely transparent. We include additional information into this static view by blending-in additional annotations.

6 Results

We have integrated the focus of attention functionality as a plugin into VolumeShop [4]. This system allows easy prototyping with the possibility of using a lot of existing functionality. We have extended the information about the dataset, which is stored in an XML file, by information on the volume up-vector and on the volume front-vector. After the viewpoint estimation, for each object a set of characteristic viewpoints is saved into the XML structure as well as the *globally* characteristic views when no specific object is in focus. Visibility computation for each object is the most time-consuming part of the pipeline and takes about few minutes, because the ray-casting has to be performed for a large number of viewpoints. As this is a pre-processing step that is considerably shorter than object specification by segmentation or setting-up proper visual representations, this is not a real issue. Interaction is done on-the-fly and additional viewpoint location computations as well as importance-driven modifications of visual representations do not take any noticeable time and the performance is equal to standard multi-volume rendering implemented in VolumeShop [4].

Focus of attention has been applied on three different data sets. The human hand and torso (Figures 1 and 2) show objects that are inside the data set. In this case the visibility computation used cut-away views to identify the best visibility. In case of the stag beetle data set (Figure 6), only outer parts have been selected so the option for cut-away visibility calculation was not necessary. In this figure sample images have been taken from re-focusing from thorax object to the legs. Between the fourth and fifth image the contextual viewpoint has been reached and the focus switched to legs.

The concept of importance-driven focusing is best demonstrated by the accompanying

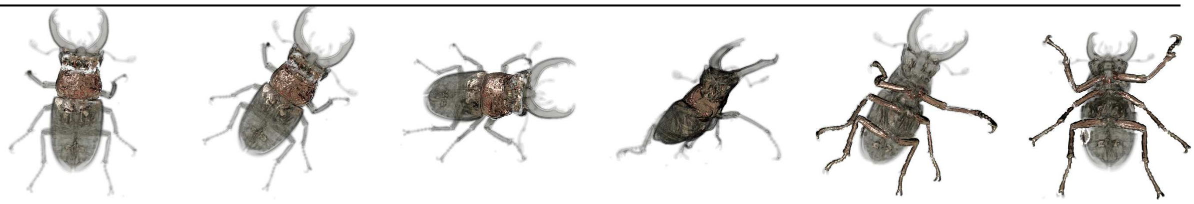

Figure 6: Stag beetle data set: re-focusing from thorax object to the legs).

video. Further information is available at http://www.cg.tuwien.ac.at/research/vis/exvisation/idf/.

7 Summary and Conclusions

In this paper we have proposed the concept for importance-driven focus of attention. We have discussed the necessary pre-processing steps before a visual inspection puts the focus of attention on interesting objects. One of these steps is localization of viewpoints that show characteristics of an object in the best way. We use a new method for viewpoint selection for volume data using viewpoint mutual information that works very good for segmented volumetric data classified by importance.

We have shown possibilities how to realize focus of attention for a visual inspection of volumetric data with added information such as varying visual representations, optimal viewpoints for objects and the entire volume, up-vector of the volume, and auxillary textual information.

We have discussed aspects of a visually pleasing re-focusing from one object of interest to another. This includes the selection of viewpoints, design of a path for the viewpoint and also changes in the visual representation. *Browsing* through pre-selected structures gives a good overview on the information content of the underlying data.

8 Acknowledgments

The work presented in this publication is carried out as part of the ex**vis**ation project (`www.cg.tuwien.ac.at/research/vis/exvisation`) supported by the Austrian Science Fund (FWF) under grant no. P18322.

This project has been funded in part with grant numbers TIN2004-07451-C03-01, FIT-350101-2004-15 of the Spanish Government and IST-2-004363 (GameTools: Advanced Tools for Developing Highly Realistic Computer Games) from the VIth European Framework.

The stag beetle from Georg Glaeser, Vienna University of Applied Arts, Austria, was scanned with an industrial CT by Johannes Kastner, Wels College of Engineering, Austria, and Meister Eduard Gröller, Vienna University of Technology, Austria. The *Monster Study* human torso and the human hand data sets are courtesy of Tiani Medgraph.

The authors would like to thank Stefan Bruckner and Peter Rautek for fruitful discussions.

References

[1] C. Andújar, P. P. Vázquez, and M. Fairén. Wayfinder: guided tours through complex walkthrough models. *Proceedings of Eurographics '04*, pages 499–508, 2004.

[2] P. Beziér. *Emploi des machines a commande num4rique*. Masson et Cie., Paris, 1970. Translated by R. Forrest and A. Pankhurst as Numerical Control – Mathematical and Applications, John Wiley and Sons Ltd., London, 1972.

[3] U. Bordoloi and H.-W. Shen. View selection for volume rendering. In *Proceedings of IEEE Visualization'05*, pages 487–494, 2005.

[4] S. Bruckner and M. E. Gröller. VolumeShop: An interactive system for direct volume illustration. In *Proceedings of IEEE Visualization'05*, pages 671–678, 2005. http://www.cg.tuwien.ac.at/volumeshop/.

[5] Digital morphology web site, http://www.digimorph.org/, 2006.

[6] B. Gooch, E. Reinhard, C. Moulding, and P. Shirley. Artistic composition for image creation. In *Proceedings of Eurographics Symposium on Rendering '01*, pages 83–88, 2001.

[7] G. Kindlmann and J. Durkin. Semi-automatic generation of transfer functions for direct volume rendering. In *Proceedings of IEEE VolVis '98*, pages 79–86, 1998.

[8] R. Kosara, S. Miksch, and H. Hauser. Semantic depth of field. In *Proceedings of IEEE InfoVis '01*, pages 97–104, 2001.

[9] R. Kosara, S. Miksch, H. Hauser, J. Schrammel, V. Giller, and M. Tscheligi. Useful properties of semantic depth of field for better F+C visualization. In *Proceedings of VisSym '02*, pages 205–210, 2002.

[10] C. H. Lee, Amitabh Varshney, and David W. Jacobs. Mesh saliency. In *Proceedings of ACM SIGGRAPH '05*, pages 659–666, 2005.

[11] J. Marks, B. Andalman, P. A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J. Seims, and S. Shieber. Design Galleries: a general approach to setting parameters for computer graphics and animation. In *Proceedings of ACM SIGGRAPH '97*, pages 389–400, 1997.

[12] M. Sbert, D. Plemenos, M. Feixas, and F. González. Viewpoint quality: Measures and applications. In *Proceedings of Computational Aesthetics '05*, pages 185–192, 2005.

[13] N. Svakhine, D. Ebert, and D. Stredney. Illustration motifs for effective medical volume illustration. *IEEE Computer Graphics and Applications*, 25(3):31–39, 2005.

[14] S. Takahashi, I. Fujishiro, Y. Takeshima, and T. Nishita. A feature-driven approach to locating optimal viewpoints for volume. In *Proceedings of IEEE Visualization'05*, pages 495–502, 2005.

[15] J. van Wijk. The value of visualization. In *Proceedings of IEEE Visualization'05*, pages 567–574, 2005.

[16] J. van Wijk and W. Nuij. Smooth and efficient zooming and panning. In *Proceedings of IEEE InfoVis '03*, pages 15–22, 2003.

[17] J. van Wijk and W. Nuij. A model for smooth viewing and navigation of large 2d information spaces. *IEEE Transactions on Visualization and Computer Graphics*, 10(4):447–458, 2004.

[18] P. P. Vázquez, M. Feixas, M. Sbert, and W. Heidrich. Viewpoint selection using viewpoint entropy. In *Proceedings of VMV '01*, pages 273–280, 2001.

[19] P. P. Vázquez, M. Feixas, M. Sbert, and W. Heidrich. Automatic view selection using viewpoint entropy and its application to image-based modeling. *Computer Graphics Forum*, 22(5):689–700, 2003.

[20] P. P. Vázquez, M. Feixas, M. Sbert, and A. Llobet. A new tool for obtaining good views for molecules. In *Proceedings of VisSym '02*, pages 183–188, 2002.

[21] I. Viola, M. E. Gröller, K. Bühler, M. Hadwiger, B. Preim, D. Ebert, M.C. Sousa, and D. Stredney. Illustrative visualization. Visualization 2005 Tutorial, 2005. http://www.cg.tuwien.ac.at/research/vis/exvisation/vis-tutorial/.

[22] I. Viola, A. Kanitsar, and M. E. Gröller. Importance-driven volume rendering. In *Proceedings of IEEE Visualization'04*, pages 139–145, 2004.

[23] I. Viola, A. Kanitsar, and M. E. Gröller. Importance-driven feature enhancement in volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):408–418, 2005.

[24] Volume data set repository, Vienna University of Technology, http://www.cg.tuwien.ac.at/research/vis/datasets/, 2006.

Appendix

Information-Theoretic Measures:
Let $\mathscr{X}$ be a finite set, let X be a random variable taking values x in $\mathscr{X}$ with distribution $p(x) = Pr[X = x]$. Likewise, let Y be a random variable taking values y in $\mathscr{Y}$. The *Shannon entropy* $H(X)$ of a random variable X is defined by

$$H(X) = -\sum_{x\in\mathscr{X}} p(x)\log p(x). \tag{8}$$

The Shannon entropy $H(X)$, also denoted by $H(p)$, measures the average uncertainty of 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* is defined by

$$H(Y|X) = -\sum_{x\in\mathscr{X}} p(x) \sum_{y\in\mathscr{Y}} p(y|x)\log p(y|x), \tag{9}$$

where $p(y|x) = Pr[Y = y|X = x]$ is the conditional probability. The conditional entropy $H(Y|X)$ measures the average uncertainty associated with Y if we know the outcome of X. In general, $H(Y|X) \neq H(X|Y)$, and $H(X) \geq H(X|Y) \geq 0$.

The *mutual information* between X and Y is defined by

$$\begin{aligned} I(X,Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= \sum_{x\in\mathscr{X}} p(x) \sum_{y\in\mathscr{Y}} p(y|x) \log \frac{p(y|x)}{p(y)}. \end{aligned} \quad (10)$$

The mutual information $I(X,Y)$ is a measure of the shared information between X and Y. It can be seen that $I(X,Y) = I(Y,X) \geq 0$.

The *relative entropy* or *Kullback-Leibler distance* between two probability distributions p and q is defined as

$$KL(p|q) = \sum_{x\in\mathscr{X}} p(x) \log \frac{p(x)}{q(x)}, \quad (11)$$

where, from continuity, we use the convention that $0\log 0 = 0$, $p(x)\log \frac{p(x)}{0} = \infty$ if $p(x) > 0$ and $0\log \frac{0}{0} = 0$. The relative entropy $KL(p|q)$ is a measure of the inefficiency of assuming that the distribution is q when the true distribution is p.

The *Jensen-Shannon divergence* between n probability distributions $p_1, p_2, \ldots, p_n$, with their corresponding weights $\pi_1, \pi_2, \ldots, \pi_n$ fulfilling $\sum_{i=1}^{n} \pi_i = 1$, is defined by

$$JS(p_1, p_2, \ldots, p_n) = H(\sum_{i=1}^{n} \pi_i p_i) - \sum_{i=1}^{n} \pi_i H(p_i). \quad (12)$$

The Jensen-Shannon divergence measures how *far* are the probabilities p_i from their mixture $\sum_{i=1}^{n} \pi_i p_i$. It equals zero if and only if all the p_i are equal.

Information-Theoretic Viewpoint Quality Measures for Polygonal Data:

Viewpoint entropy measure based on Shannon entropy (Equation 8) is defined as

$$H_v = -\sum_{i=0}^{N_f} \frac{a_i}{a_t} \log \frac{a_i}{a_t}, \quad (13)$$

where N_f is the number of polygons of the scene, a_i is the projected area of polygon i over the sphere of directions centered at viewpoint v, a_0 represents the projected area of background in open scenes, and $a_t = \sum_{i=0}^{N_f} a_i$ is the total area of the sphere. The maximum entropy is obtained when a certain viewpoint can see all the polygons with the same projected area a_i.

Viewpoint measure based on Kullback-Leibler distance (Equation 11) is defined by

$$KL_v = \sum_{i=1}^{N_f} \frac{a_i}{a_t} \log \frac{\frac{a_i}{a_t}}{\frac{A_i}{A_T}}, \quad (14)$$

where a_i is the projected area of polygon i, $a_t = \sum_{i=1}^{N_f} a_i$, A_i is the actual area of polygon i and $A_T = \sum_{i=1}^{N_f} A_i$ is the total area of the scene or object.