Similarity-Based Exploded Views

Marc Ruiz[1], Ivan Viola[2], Imma Boada[1], Stefan Bruckner[3], Miquel Feixas[1],
and Mateu Sbert[1]

[1] Graphics and Imaging Laboratory, University of Girona, Spain
[2] Department of Informatics, University of Bergen, Norway
[3] Institute of Computer Graphics and Algorithms,
Vienna University of Technology, Austria

Abstract. Exploded views are often used in illustration to overcome the problem of occlusion when depicting complex structures. In this paper, we propose a volume visualization technique inspired by exploded views that partitions the volume into a number of parallel slabs and shows them apart from each other. The thickness of slabs is driven by the similarity between partitions. We use an information-theoretic technique for the generation of exploded views. First, the algorithm identifies the viewpoint which gives the most structured view of the data. Then, the partition of the volume into the most informative slabs for exploding is obtained using two complementary similarity-based strategies. The number of slabs and the similarity parameter are freely adjustable by the user.

1 Introduction

Volume visualization aims at gaining insight into volumetric data using interactive graphics and imaging techniques. Current volume data sets generated by scientific domains contain large amounts of data of complex structures. Effective visualization of such data sets that clearly shows all contained structures is challenging.

Illustrative visualization enhances the expressiveness of volume rendering by applying hand-crafted illustrative techniques. Cut-aways, exploded views or high-level abstraction strategies, amongst others, are used to reveal insights and represent essential structures of the volume in a clear way while less important details are subjugated. To employ these techniques, certain controlling mechanisms based on data or higher semantical levels (e.g. segmentation into objects from the domain perspective and the assigning of object importance based on the given domain scenario) are required. These mechanisms vary from fully interactive steered by user (e.g. voxel-by-voxel segmentation) to fully automatic techniques (e.g. shape analysis of the acquired data based on higher-order derivatives). To explore unclassified data sets, automatic controlling mechanisms for steering expressive visualization are useful, and possibly can be combined with interactive techniques that *fine-tune* the first automatic *educated guess*.

Our interest is focused on exploded views, which partition the volume into different parts that are displaced away from each other as if there had been a small

controlled explosion emanating from the focus of interest. Exploded views enable to see details of otherwise overlapping structures, exploiting the observer's understanding of the original spatial arrangement. In this paper, a new partitioning approach for automatic generation of exploded views is presented. This method divides the data set into a set of slabs defined by parallel planes, combining in this way the advantages of 2D and 3D views. While 3D visualization provides a global view of the entire model, the 2D cross sectional views reveal insights. To partition the volume, two alternative strategies are proposed. The first one starts with the entire volume and partitions it recursively guided by a *maximum dissimilarity* criterion. The second one considers initially all individual slices and groups them together according to a *similarity* criterion. In both cases, the controlling mechanism is the similarity value that is computed automatically using information-theoretic measures. The only necessary interaction of the user with the data is a single threshold parameter which determines when the partitioning (or grouping) has to stop. An important advantage of this approach is that no a-priori information or pre-processing of the data is required. This is suitable, especially, for computer-guided exploration of histology volume data.

2   Related Work

The main limiting factor when exploring volume data is the occlusion between structures. For complex volumetric data sets it is difficult to achieve a visual representation that not only shows all the internal structures but also preserves the global representation of the model. To enhance volume data interpretation Rheingans and Ebert [1] introduced the volume illustration approach, combining the familiarity of a physics-approximated illumination model with the ability to enhance important features using non-photorealistic rendering techniques. Volume illustration techniques enhance the perception of structure, shape, orientation, and depth relationships in a volume model. Although they cannot totally solve the occlusion problem, the good performance of these techniques led to the development of new volume rendering methods.

Clipping away or removing away parts of the data to eliminate occlusion is a well-known and extensively used approach. The loss of context due to removed parts is the main limiting factor of such a technique. To overcome this limitation, strategies with more complex clipping geometry have been proposed. Wang et al. [2] introduced volume sculpting as a flexible approach to explore data. Weiskopf et al. [3] proposed several interactive clipping techniques that are capable of using complex clip geometries. Konrad-Verse et al. [4] described a method which is based on a deformable cutting plane for virtual resection. Viola et al. [5] presented an importance-driven approach capable of enhancing important features while preserving the necessary context by generating cut-away views and ghosted images from volumetric data. Bruckner et al. [6] proposed an alternative to conventional clipping techniques in order to avoid loss of context. Their context-preserving volume rendering model uses a function of shading intensity, gradient magnitude, distance to the eye point, and previously accumulated opacity to selectively reduce the opacity in less important data regions.

Exploded views and deformations are a common strategy for communicating the structure of complex 3D objects that are composed of many subparts. Deformation metaphors for browsing structures in volumetric data were introduced in volume visualization by McGuffin et al. [7]. They presented an approach for volume exploration based on deformations that allows the users to cut into and open up, spread apart, or peel-away layers of the volume while still retaining the surrounding context. The explosion of the parts is set manually. Bruckner et al. [6] went one step further by automating the explosion. Their method uses a continuous degree-of-interest function to distinguish between focus and context and is capable of re-arranging the parts dynamically based on the viewpoint. In these techniques, a priori knowledge of the volume data to define the layers or to set the focus of interest is assumed – the data has been explicitly partitioned by the user. In contrast, our approach automatically partitions the volume based on characteristics of the data.

Moreover, good viewpoint selection is also crucial for an effective focus of attention [5]. Different information-theoretic measures for viewpoint evaluation have been presented. Vàzquez et al. [8] have introduced the viewpoint entropy as a measure for viewpoint quality evaluation, where the best viewpoint is defined as the one that has maximum entropy. Designed for polygonal data, this measure has been extended to volumetric scalar data [9,10]. Viola et al. [11] have presented the viewpoint mutual information from the definition of an information channel between a set of viewpoints and a set of objects of a volumetric data set. This measure provides representative views and is very robust with respect to the resolution of the model.

3   Similarity-Steered Visualization

To automatically obtain the partitioning planes for the exploded views, we propose a two-step process. First, we select the view of the model along which the organs or components will be better separated. This view is called the *most structured view* of the model. Second, we calculate the partitions of the model along the most structured view. Such partitions will be obtained using two complementary approaches: a top-down strategy that divides the model according to the maximum information gain and a bottom-up method that joins the slices according to a similarity criterion. Then, the explosion of the model is visualized in the interactive system VolumeShop [12]. The two steps of the method are described below.

1. **Selection of splitting axis**
   The goal of this step is to obtain the most structured view of the model. To reach this objective a viewpoint measure able to capture the structure of the volumetric dataset along any view axis is used. In information theory, *entropy rate* is defined as a measure of the irreducible randomness of an object or the degree of unpredictability of a sequence of values. Since a high randomness corresponds to a low structure and vice versa, we can use the entropy rate
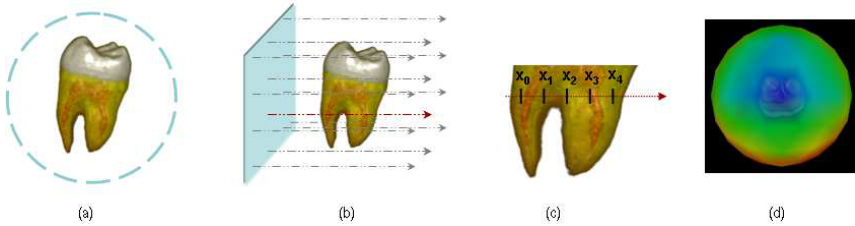
Fig. 1. Main steps of the selection of the most structured view. (a) Sphere of viewpoints, (b) sampling process for one viewpoint, (c) samples considered for the entropy rate computation, and (d) colored viewpoint sphere (using a thermic scale, from blue to red) representing the values of the viewpoint entropy rate.

to quantify the degree of structure or predictability of a model. We proceed as illustrated in Figure 1. First of all, the model is centered in a viewpoint sphere built from the recursive discretisation of an icosahedron (Figure 1(a)). Then, for each viewpoint the entropy rate is computed as described in Section 4 (Figure 1(b,c)). Finally, we identify the lowest entropy rate value which corresponds to the most structured view of the model (Figure 1(d)). This direction is used as axis to which similarity-based partitioning planes are perpendicular to.

2. **Volume partitioning**

   This task consists of selecting the optimal partitions of the model from the most structured view. To carry out this process two different strategies are presented:

   (a) **Top-down approach.** Initially, the entire volume is considered and partitioning planes are taken perpendicular to the most structured view (Figure 2(i.a)). To divide the dataset into different parts, we use a greedy algorithm which successively selects the partition that provides us with the maximum gain of information. According to the information bottleneck method [13,14], the information gain can be calculated using the Jensen-Shannon divergence between two parts of the model (Figure 2(i.b)). This measure can be interpreted as the degree of dissimilarity between the parts and attempts to divide the model into homogeneous regions (Figure 2(i.c)). A more detailed description of this approach is given in Section 5.1.

   (b) **Bottom-up approach.** All the slices of the volume, perpendicular to the most structured view, are considered as the initial slabs (Figure 2(ii.a)). Neighboring slabs are iteratively grouped (Figure 2(ii.b)) when mutual information between them is higher than a given threshold (Figure 2(ii.c)). Dealing with similarity between slabs instead of individual slices, we avoid an incorrect grouping, for instance, due to smooth changes along many consecutive slices. The grouping process is further described in Section 5.2.
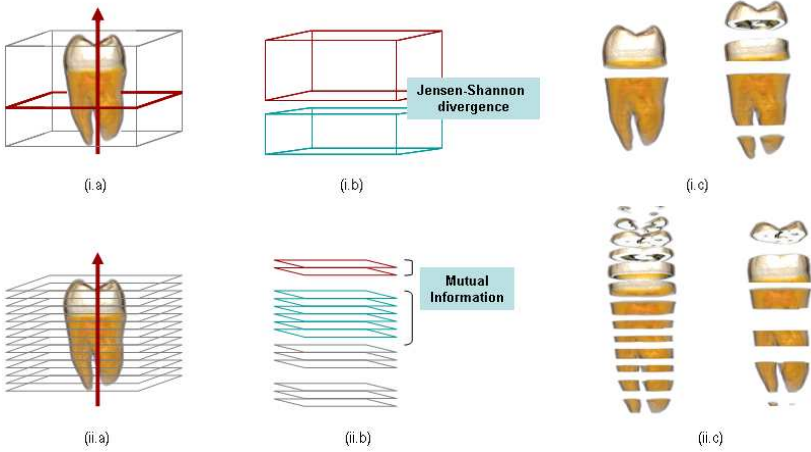
Fig. 2. (i) Top-down volume partition: (i.a) partitioning planes are taken perpendicular to the most structured view, (i.b) dissimilarity between subvolumes is given by the Jensen-Shannon divergence, and (i.c) examples showing two different partitions. (ii) Bottom-up volume partition: (ii.a) slices are taken perpendicular to the most structured view direction, (ii.b) similarity between slices or slabs is computed using mutual information, and (ii.c) two examples resulting from the grouping process.

4   Selection of Structured Views

To quantify the degree of structure of a volumetric data set along a given viewing direction, we estimate the entropy rate of the sequence of values (intensities) obtained by casting a bundle of parallel lines along that direction. These lines act as probes to sample the intensity of the model. The view with the lowest entropy rate will correspond to the most structured view.

The definitions of both Shannon entropy and entropy rate [15] are now reviewed. The notation used is inspired by the work of Feldman and Crutchfield [16]. Let $\mathcal{X}$ be a finite set and X a random variable taking values x in $\mathcal{X}$ with probability distribution $p(x) = Pr[X = x]$. The *Shannon entropy* $H(X)$ of a random variable X is defined by

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

The Shannon entropy measures the average uncertainty of random variable X. If the logarithms are taken in base 2, entropy is expressed in bits.

Given a sequence $X_1 X_2 \ldots$ of random variables X_i taking values in $\mathcal{X}$, a block of L consecutive random variables is denoted by $X^L = X_1 \ldots X_L$. The probability that the particular L-block x^L occurs is denoted by joint probability $p(x^L) = p(x_{i+1}, \ldots, x_{i+L})$. The *joint entropy* of a block of L consecutive symbols or *L-block entropy* is defined by

$$H(X^L) = - \sum_{x^L \in \mathcal{X}^L} p(x^L) \log p(x^L), \tag{2}$$

where the sum runs over all possible L-blocks.

The *entropy rate* or *entropy density* is defined by

$$h = \lim_{L \to \infty} \frac{H(X^L)}{L} \tag{3}$$

and measures the average amount of information per symbol x [15].

It can also be rewritten as

$$h = \lim_{L \to \infty} (H(X^L) - H(X^{L-1})). \tag{4}$$

The entropy rate of a sequence of symbols is a measure of its uncertainty, randomness or unpredictability. The entropy rate is also a measure of the compressibility of a sequence: the higher the uncertainty, the lower the compressibility. For instance, in a text, if there are strong correlations between letters (or words), knowledge of all previous letters (or words) will significantly decrease our uncertainty about the next one [16].

How to compute the entropy rate for a given viewpoint is now shown. Consider the scheme in Figure 1. For each viewpoint, a sequence of samples (intensity values) to compute the measure is obtained performing a ray casting from a plane centered at the viewpoint. We proceed as follows:

- From the plane at each viewpoint, parallel rays with a regular horizontal and vertical spacing x are cast. Along the ray within the volume, equidistant samples at distance y are taken.
- To build the two joint histograms of X^L and X^{L-1} required for the entropy rate computation, we take into account all possible groups of consecutive samples of length L and $L-1$, respectively. For example, with the samples shown in Figure 1(c), we can form three blocks of length 3 ($x_0x_1x_2$, $x_1x_2x_3$ and $x_2x_3x_4$) and four of length 2 (x_0x_1, x_1x_2, x_2x_3 and x_3x_4) for X^3 and X^2 histograms, respectively.
- From the joint histograms of X^L and X^{L-1}, the joint probability distributions $p(x^L)$ and $p(x^{L-1})$ are estimated and then the joint entropies $H(X^L)$ and $H(X^{L-1})$ are calculated.
- Due to the potentially high dimensionality of the histograms and, consequently, the high number of components, a trade-off between the number of symbols (intensities) and the length L of the blocks has to be considered. Note that the size (number of entries) of the highest histogram is $O(N^L)$, where N is the number of different property values and L is the length of the blocks. Usually voxel models have property values of 8 bits or more, so this problem is untreatable even with short blocks. As entropy rate is a limit quantity (4), its computation would require an infinite number of elements and blocks infinitely long. It has to be approximated using a block of finite length. From the two possible approximations coming from (4) and (4), we
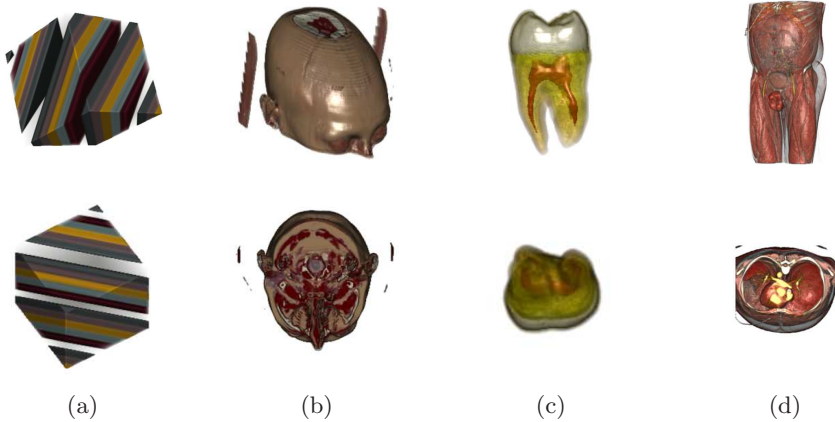
| (a) | (b) | (c) | (d) |

Fig. 3. Different volumes (first row) and their corresponding most structured views (second row). From left to right: (a) a synthetic model, (b) a CT-scan of a patient with hemorrhage, (c) a CT-scan of a tooth and (d) a CT-scan of the human body.

have selected the last one because it approximates more accurately the entropy rate for low values of L. In our experiments, we have taken $L = 3$. We have also reduced the number of symbols of $\mathcal{X}$ intensity bins in the histogram to 32, rescaling the intensity bins.

The strategy for the selection of the most structured view has been applied to different volume data sets. The obtained results are illustrated in Figure 3 where the first row represents the original model and the second row the most structured view. From left to right, the proposed models correspond to: (a) a synthetic model, created considering six different materials (each one represented with a different color) which follow a diagonal distribution through the volume, (b) a CT-scan of a patient with an hemorrhage, (c) a CT-scan of a tooth and (d) a CT-scan of the human body. Observe how the best views show the maximum of structure in the model. This is specially noticeable in the phantom model (Figure 3(a)) where the different regions have an inclination relative to the cube axis.

5   Evaluating Similarity

To obtain the optimal partitions for the explosion of a 3D data set, two different strategies are presented. First, we analyze a top-down approach which partitions the model using a criterion of maximum gain of information. Second, we study a bottom-up strategy that groups the slices according to a similarity measure between them.

5.1   Model Partitioning

Once the most structured direction of the model has been selected, a sequence of perpendicular partitions in that direction can be obtained using a simple greedy

algorithm. This is a top-down hierarchical application of the information bottle-neck method [13,14] which permits us to measure the gain of information when a model is divided into different slabs. This *gain of information* is computed using the Jensen-Shannon divergence.

The *Jensen-Shannon divergence* [17] between probability distributions p_1, $p_2, \ldots, p_N$ with prior probabilities or weights $\pi_1, \pi_2, \ldots, \pi_N$ is defined by

$$JS(\pi_1, \pi_2, \ldots, \pi_N; p_1, p_2, \ldots, p_N) \equiv H\left(\sum_{i=1}^{N} \pi_i p_i\right) - \sum_{i=1}^{N} \pi_i H(p_i), \qquad (5)$$

where $\sum_{i=1}^{N} \pi_i = 1$. The JS-divergence measures how far the probabilities p_i are from their likely joint source $\sum_{i=1}^{N} \pi_i p_i$ and equals zero if and only if all the p_i are equal. From [14], it can be seen that the gain in information when a dataset is divided into two slabs is given by

$$\Delta I = \frac{v_1 + v_2}{v_T} JS(\frac{v_1}{v_1 + v_2}, \frac{v_2}{v_1 + v_2}; p_1, p_2), \qquad (6)$$

where v_1 and v_2 are, respectively, the volumes of slabs 1 and 2, v_T is the total volume of the 3D dataset, p_1 and p_2 are, respectively, the normalized intensity histograms of slabs 1 and 2, and $JS(\frac{v_1}{v_1+v_2}, \frac{v_2}{v_1+v_2}; p_1, p_2)$ is the Jensen-Shannon divergence between p_1 and p_2 with the corresponding weights $\frac{v_1}{v_1+v_2}$ and $\frac{v_2}{v_1+v_2}$.

The gain of information when a model is divided into two parts is given by the dissimilarity between them (measured by JS) weighted by their relative volume. Note that a slab highly structured along a given direction will have all possible partitions very similar and thus will not need to be partitioned.

5.2   Slice Grouping

Given a viewing direction, the slices perpendicular to it can be grouped using a similarity measure. In this paper, the normalized mutual information is used to quantify the degree of similarity between individual slices or groups of adjacent slices (slabs). In medical imaging, many successful automatic image registration methods are based on the maximization of mutual information. This method, introduced by Viola [18] and Maes et al. [19], is based on the conjecture that the correct registration corresponds to the maximum mutual information between the overlap areas of the two images. Later, Studholme et al. [20] proposed to use the normalized mutual information as it is more robust than solely mutual information due to its greater independence of the overlap area.

Let X and Y be two random variables taking values x and y in finite sets $\mathcal{X}$ and $\mathcal{Y}$ with probability distributions $p(x) = Pr[X = x]$ and $p(y) = Pr[Y = y]$, respectively. The *mutual information* between X and Y is defined by

$$MI(X,Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \qquad (7)$$

where $p(x,y) = Pr[X = x, Y = y]$ is the joint probability. MI is a measure of the shared information or the degree of dependence between X and Y. MI is

zero only if the two random variables are strictly independent. The *normalized mutual information* is defined by

$$NMI(X,Y) = \frac{MI(X,Y)}{H(X,Y)}, \tag{8}$$

where $H(X,Y)$ is the joint entropy of X and Y. NMI takes values in the range [0,1].

An explanation is now given to compute the NMI measure between slices and the algorithm to group them. Given two slices A and B from the volume dataset, with associated random variables X and Y, the joint probability distribution $p(x,y)$ can be estimated by simple normalization of the joint histogram $h(x,y)$ of both slices. This is obtained from the intensities of each pair (a,b) of corresponding voxels, where $a \in A$ and $b \in B$. Once the joint histogram has been calculated, the joint probability distribution and the marginal probability distributions of X and Y can be estimated: $p(x,y) = \frac{h(x,y)}{\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} h(x,y)}$, $p(x) = \sum_{y \in \mathcal{Y}} h(x,y)$ and $p(y) = \sum_{x \in \mathcal{X}} h(x,y)$. The similarity measure NMI is then evaluated.

The similarity between two slices can be extended to the similarity between two slabs $\widehat{A} = \{A_1, \ldots, A_n\}$ and $\widehat{B} = \{B_1, \ldots, B_m\}$. The random variables $\widehat{X}$ and $\widehat{Y}$, associated with both slabs, represent the grouping of a set of random variables $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$, respectively. Their joint frequency histogram is obtained from the intensities of each pair of corresponding voxels (a_i, b_j), where $a_i \in A_i$ and $b_j \in B_j$ $\forall i,j$. As mentioned above, the joint and marginal probability distributions can be estimated and thus the NMI measure is obtained.

Given the similarity measure NMI, the algorithm proceeds by joining the two adjacent slabs with maximum similarity. This process stops when the similarity between them is above a user-defined threshold or a number of slabs has been reached. At the beginning, every slab consists of only one slice. Then, the most similar slabs are progressively joined. To group n slices, the algorithm proceeds as follows:

- Assign n slabs such that each slab contains exactly one slice.
- Compute NMI for each pair of consecutive slabs.
- Find the two closest consecutive slabs i and $i+1$ (with maximum NMI). If the similarity between them is higher than the given threshold, then create a new slab $\widehat{i}$ by combining i and $i+1$ and recalculate NMI for the neighboring slabs of $\widehat{i}$. This step stops when the similarity between each pair of consecutive slabs is lower than a fixed threshold or a number of slabs is achieved.

5.3   Results

These proposed approaches have been implemented and integrated into the VolumeShop framework [12]. To test the methods, different synthetic and real data sets have been considered. In all the tests a sphere of 42 viewpoints has been
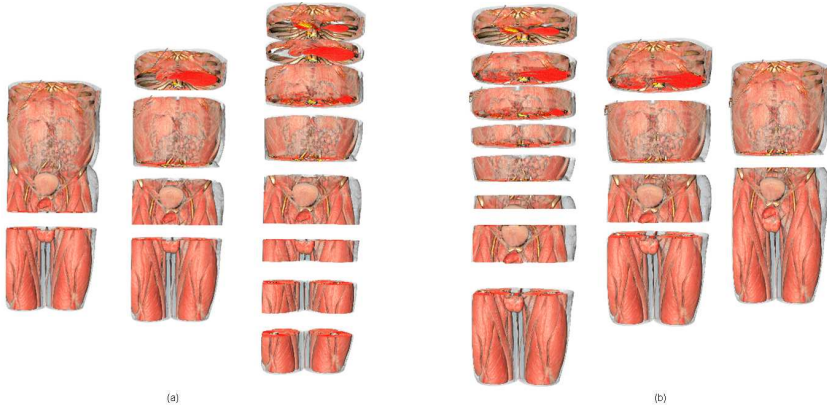
Fig. 4. (a) The volume partitioning and (b) the slice grouping approaches applied to a CT scan of the human body
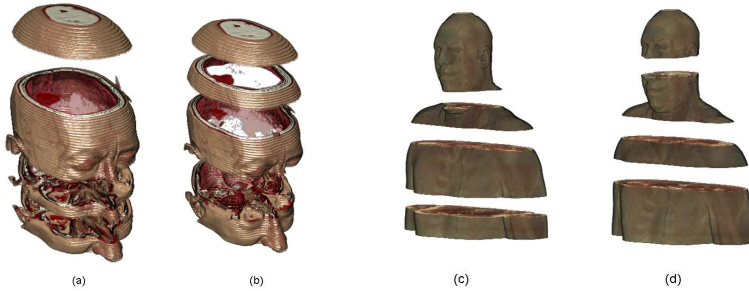

Fig. 5. (a) The volume partitioning and (b) the slice grouping approaches applied to a CT-scan of a patient with a brain hemorrhage. An histological data model decomposed with the (c) volume partitioning and (d) slice grouping methods.

used and the stopping criterion has been fixed by the number of slabs entered by the user.

For the first tests, the CT scan of the human body of Figure 4 has been used. The results obtained with the model partitioning approach for 2, 4 and 8 partitions are illustrated in Figure 4(a). In Figure 4(b) the partitions obtained with the slice grouping approach using the same user parameters are shown. In Figures 5(a) and (b) we illustrate the results obtained by applying the volume partitioning and the slice grouping approaches on a CT scan of a patient with a brain hemorrhage. Observe that the damaged region is located in the second slab from top to bottom.

In Figures 5(c) and (d), the results obtained with the volume partitioning and the slice grouping approaches applied to an histologic data model are shown. It is important to emphasize that these techniques have been applied without prior pre-processing. Time cost for computing the most structured view, and for volume partitioning and slice grouping are given in Table 1. Times are

Table 1. Time cost (seconds) for computing the most structured view, and for volume partitioning (P) and slice grouping (G) in 2, 4, and 8 slabs, respectively

Volume	Size	Best View	P(2)	P(4)	P(8)	G(8)	G(4)	G(2)
Human body	$256 \times 256 \times 415$	111.0	0.4	0.9	2.0	28.9	36.2	70.5
Hemorrhage	$512 \times 512 \times 45$	55.5	0.2	0.6	1.2	6.3	10.8	14.4
Histological data	$587 \times 342 \times 499$	722.6	2.3	5.8	12.6	214.0	324.7	525.0

given for a CPU Intel(R) Core(TM)2 Quad CPU Q6600 at 2.40GHz with 2 GB of memory. Several video sequences are available as supplementary material in `http://www.gametools.org/smartgraphics/`.

6   Conclusions

New partitioning techniques for volumetric data decomposition and visualization using exploded views have been introduced. These techniques use an information-theoretic two-step approach to automatically partition the model. First, the view with the highest structure is identified and, then, the model is divided along this view following two alternative similarity-based methods. The presented techniques provide us an efficient tool for volume data exploration without neither a priori knowledge nor pre-processing of the data. In our future work we will study how many slices should be presented and whether additional information about the similarity distances can help the user for an optimal understanding of the volume. Also, the trade-off between quality and cost for the different parameters involved in the determination of most structured views (distances between rays, subsequences of samples on those rays, sub-sampling of intensity values) will be investigated.

References

1. Rheingans, P., Ebert, D.: Volume illustration: Nonphotorealistic rendering of volume models. IEEE Trans. on Visualization and Computer Graphics 7(3), 253–264 (2001)
2. Wang, S.W., Kaufman, A.E.: Volume sculpting. In: Proc. the Symposium on Interactive 3D Graphics, pp. 151–156 (1995)
3. Weiskopf, D., Engel, K., Ertl, T.: Interactive clipping techniques for texture-based volume visualization and volume shading. IEEE Trans. on Visualization and Computer Graphics 9(3), 298–312 (2003)

4. Konrad-Verse, O., Preim, B., Littmann, A.: Virtual resection with a deformable cutting plane. In: Proc. of Simulation und Visualisierung, pp. 203–214 (2004)
5. Viola, I., Kanitsar, A., Gröller, M.E.: Importance-driven feature enhancement in volume visualization. IEEE Trans. on Visualization and Computer Graphics 11(4), 408–418 (2005)
6. Bruckner, S., Grimm, S., Kanitsar, A., Gröller, M.E.: Illustrative Context-Preserving Exploration of Volume Data. IEEE Trans. on Visualization and Computer Graphics 12(6), 253–264 (2006)
7. McGuffin, M., Tancau, L., Balakrishnan, R.: Using deformations for browsing volumetric data. In: Proc. IEEE Visualization, pp. 401–408 (2003)
8. Vázquez, P.P., Feixas, M., Sbert, M., Heidrich, W.: Viewpoint selection using viewpoint entropy. In: Proc. Vision, Modeling, and Visualization 2001, pp. 273–280 (2001)
9. Takahashi, S., Fujishiro, I., Takeshima, Y., Nishita, T.: A feature-driven approach to locating optimal viewpoints for volume visualization. In: IEEE Visualization 2005, pp. 495–502 (2005)
10. Bordoloi, U.D., Shen, H.W.: Viewpoint evaluation for volume rendering. In: IEEE Visualization 2005, pp. 487–494 (2005)
11. Viola, I., Feixas, M., Sbert, M., Gröller, M.E.: Importance-driven focus of attention. IEEE Trans. on Visualization and Computer Graphics 12(5), 933–940 (2006)
12. Bruckner, S., Gröller, M.E.: Volumeshop: An interactive system for direct volume illustration. In: Proc. IEEE Visualization 2005, pp. 671–678 (2005)
13. Tishby, N., Pereira, F.C., Bialek, W.: The information bottleneck method. In: Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing, pp. 368–377 (1999)
14. Slonim, N., Tishby, N.: Agglomerative information bottleneck. In: NIPS, pp. 617–623 (1999)
15. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley Series in Telecommunications (1991)
16. Feldman, D., Crutchfield, J.: Structural information in two-dimensional patterns: Entropy convergence and excess entropy. Physical Review E 67 (2003)
17. Burbea, J., Rao, C.R.: On the convexity of some divergence measures based on entropy functions. IEEE Trans. on Information Theory 28(3), 489–495 (1982)
18. Viola, P.A.: Alignment by Maximization of Mutual Information. PhD thesis, Massachusetts Institute of Technology, Massachusetts (MA), USA (1995)
19. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. on Medical Imaging 16(2), 187–198 (1997)
20. Studholme, C.: Measures of 3D Medical Image Alignment. PhD thesis, University of London, London, UK (1997)