Efficient complementary viewpoint selection in volume rendering

Sergi Grau

Dept. Llenguatges i
Sistemes Informatics,
Polytechnic University of
Catalonia.

sgrau@lsi.upc.edu

Anna Puig, Sergio
Escalera, Maria Salamó

Dept. Matemàtica
Aplicada i Anàlisi,
Universitat de Barcelona.

{anna,sergio,maria}@maia.ub.es

Oscar Amorós

Dept. Matemàtica
Aplicada i Anàlisi,
Universitat de Barcelona.

morousg@gmail.com

ABSTRACT

A major goal of visualization is to appropriately express knowledge of scientific data. Generally, gathering visual information contained in the volume data often requires a lot of expertise from the final user to setup the parameters of the visualization. One way of alleviating this problem is to provide the position of inner structures with different viewpoint locations to enhance the perception and construction of the mental image. To this end, traditional illustrations use two or three different views of the regions of interest. Similarly, with the aim of assisting the users to easily place a good viewpoint location, this paper proposes an automatic and interactive method that locates different complementary viewpoints from a reference camera in volume datasets. Specifically, the proposed method combines the quantity of information each camera provides for each structure and the shape similarity of the projections of the remaining viewpoints based on Dynamic Time Warping. The selected complementary viewpoints allow a better understanding of the focused structure in several applications. Thus, the user interactively receives feedback based on several viewpoints that helps him to understand the visual information. A live-user evaluation on different data sets show a good convergence to useful complementary viewpoints.

Keywords

Dual camera; Visualization; Interactive Interfaces; Dynamic Time Warping.

1 INTRODUCTION

One of the most important visualization goals is to appropriately express knowledge of scientific data. During the last few decades several methods have been published in the bibliography to gather visual information contained in the data. However, gathering visual information often requires the expertise from the final user in a difficult and tedious process in order to setup the parameters of the visualization. Some metaphors of interaction have been provided to help users in data navigation between focus and context (importance-driven, VolumeShop, exoVis, LiveSync++, ClearView).

Over the centuries, the traditional paradigm used in illustration for visual abstraction has been to enhance the most important structures into an environment or context with different painting techniques. As an example, Fig. 1 shows two traditional illustrations which use two or three different views of the region of interest to enhance the construction of the mental image. Particularly, in Fig. 1(a) the focused structure is an eye which shows two complementary viewpoints, an oblique view and a frontal view, and Fig. 1(b) illustrates complementary views of the foot.

Previously, we proposed an enhancement of the ClearView paradigm where the context's volume data

(a)   **(b)**
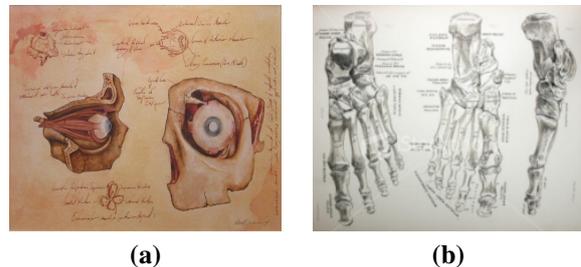
Figure 1: Illustration examples of the complementary camera viewpoint. Images (a) and (b) come from http://www.keithtuckerart.com/Illustration.html and the medical dictionary Allen's Anatomy, respectively.

is adaptively clipped around the focus region [GP09]. Apart from the focus, to select a good starting viewpoint of the focus is a difficult task as volume renderings often include a barrage of complex 3D structures that can overwhelm the user. Several methods have been published to gather visual information contained in the data. There are methods in the field of illustrative visualization [RBGV08], such as [VKG05] and [BG05], where the main goal is to develop applications that can integrate illustrations in the expert's ordinary data analysis in order to get more semantics from the data.

Once a good focus and context is obtained, it is also important to concentrate on alleviating the normally difficult abstraction process that a user should carry out to convey the desired information in the underlying data. Specifically, to select a good starting viewpoint is sometimes a tedious task for a non-experimented users. Actually, in the clinical routine prefixed views are used based on saggital, axial and coronary views, despite that they are not always the best. In the bibliography, we can find techniques that automatically locate the viewpoint according to the importance of the structure to be rendered [BS05a, TFTN05, PPVN08]. Most of them are computed in a preprocessed stage due to their high computational cost.

In this paper we propose the computation of the correlated view that shows highly complementary information of the focused structures. Also, in our proposal the user can adjust suggested cameras interactively, and our system computes the new complementary view. To this end, we consider both the visual information and the shape similarity of the projections of each sampled camera. Similarity is based on Dynamic Time Warping, ranking the projections using a weighted similarity related to the reference camera. In this manner, the proposal is able to cope with the problem of locating automatically the complementary camera that maximizes the complementary information of the structures. By its nature, the proposed algorithm is general enough to be applied in combination to state-of-the-art approaches for selecting the optimal camera. Based on a live-user evaluation, we show how the proposal efficiently computes complementary viewpoints and rapidly converges to a good solution.

The rest of the paper is structured as follows: Section 2 describes relevant previous work on viewpoint selection. Our proposal is described in Section 3. Section 4 analyzes the proposal with different data sets. Finally, Section 5 concludes the paper.

2 RELATED WORK

The cognitive process to comprehend the meaningful structures in an image varies according to the visual purpose. A single visual event is easily processed in an intuitive way. However, multiple visual events involve a more complex cognitive process. Several studies have addressed this problem. This section summarizes previous works on visual information measures used to define the best viewpoint selection.

Setting the camera to focus on the relevant structures of the model sometimes requires a lot of user expertise. Therefore, many authors [BS05a, TFTN05, VFSG06, MNTP07a, PPVN08, KBKG08, ZWD12] have addressed viewpoint optimization in direct volume rendering, extending the ideas used in surface-based scenes.

Visual information is the measure used to estimate the quality of a viewpoint according to the perception of the structures of interest of a volume data set. Then, information theory approaches, such as viewpoint entropy [BS05a], generalize the analysis of the visual information associated to a specific viewpoint. Particular cases are the approaches focused on heuristic functions [KBKG08, TFTN05, MNTP07b, CQWZ06, GP09]. These metrics are not universal and, as [PPB*05] concludes, not one descriptor performs a perfect job. Several descriptors to measure the visual information associated to a view projection have been previously described in the literature. Heuristic functions and entropy-based methods mainly consider three types of descriptors: environment-dependent [KBKG08, PPVN08], object-dependent [VFSG06, BS05a, KBKG08] and viewpoint-dependent [ZWD12].

First, environment-dependent descriptors depend on contextual parameters such as patient orientation and viewpoint history [KBKG08, PPVN08]. Object-dependent descriptors characterize the information contained in the data set, such as the importance of the selected objects [VFSG06], the noteworthiness of each voxel that contributes in the final image [BS05a], the shape of the selected object [KBKG08], and the use of heuristics for complementary views [GP09]. These type of descriptors depend upon two main requirements: a complete model of each view and the position of the camera. The process of all these models involves a high computational cost that is mostly alleviated by introducing a preprocessing stage for each view. Finally, viewpoint-dependent descriptors are based on the visibility and the location of the final projection of the object in the viewport. They measure only the information that will be effectively seen by the user. The feature most considered are the size of the projected area of the object [MNTP07b], the occlusion between objects [MNTP07b, KBKG08], object self-occlusions [CQWZ06], the important viewport areas [VFSG06] and the entropy of the image [PPVN08]. Some approaches study correlations between cameras in varying time measuring the stability between the views based on the Jense-Shanon divergence metric [BS05b]. Also, in path-views searching a Normalized Compression Distance is used [PPVN08]. They search the next best view at the greatest distance from the previous one. [CMH08] proposes a global search of optimal points in the solution space using potential fields.

Our main goal in this work is to measure the information that will be effectively seen by the user in two correlated views. The requirements are that the measure should be simple, easy to evaluate, robust to changes in the resolution of the final view, and without user intervention. Viewpoint-dependent descriptors are the

most suitable since they are focused on the final visual user's perception. They are easy to compute and they do not require a preprocess stage such as object-dependent ones. In addition, we propose to measure a new descriptor based on the shape similarity of the projected structures that complements the projected area information. We present a 2D-geometrical approach based on Dynamic Time Warping that matches distortions between the projected shapes. This method is based on implicit distance functions that are stable and robust to shape perturbations and noise. Thus, our proposal obtains a representative complementary projection using a viewpoint-dependent approach, i.e. analyzing the visual information associated to each view by an entropy-based method that combines the entropies of each projected visible structure and shape similarities.

3 LOCATING COMPLEMENTARY VIEWPOINTS

The main goal of our system is to provide an interactive exploration of volume data sets by unexperienced users. To visualise the volume, we use an interactive interface, allowing users to both understand and control aspects of the visualization process that would otherwise go unnoticed.

In order to define the location of the complementary camera of the underlying structures of the data from a reference viewpoint, v_r, we calculate the entropy of a set of sampled cameras, $H(v)$. Next, we analyze the shape similarity, *Sim*, of the viewpoint projections related to the v_r viewpoint based on Dynamic Time Warping. The weighted combination of these measures produces the final complementary view, $Dual(v_r)$.

3.1 Visual information descriptor

We use an entropy-based method to estimate the quality of a camera based on the projected area of the isosurfaces of the selected structures. Multiple focused regions are considered in the visual information estimation. We calculate the weighted sum of the viewpoint entropy's of the extracted structures. We assign higher weights to these structures with opacity transfer functions, as in [TFTN05]. Let $F = \{f_1, f_2, ..., f_f\}$ be the set of user-selected structures, or features, we define the function, $w(f_j)$, as the weight associated to the each feature, f_j. Let $H_{f_j}(v)$ be the entropy of the viewpoint, v, associated to the f_j feature. Thus, the final entropy associated to a viewpoint, v, is defined as:

$$H(v) = \sum_{j=1}^{|F|} H_{f_j}(v) \cdot w(f_j). \qquad (1)$$

The entropy $H_{f_j}(v)$ is based on view-dependent descriptors. In the bibliography, the most used are the

size of the projected area of the object [PPVN08], the occlusion between objects [KBKG08], object self-occlusions, the important viewport areas [VFSG06], and the luminance [ZWD12]. In this proposal, without loss of generality, we use as a proof of concept, the projected area of the voxels of the feature f_j.

Finding a good view assumes that the data set is centered at the origin and the camera is restricted to be at a fixed distance from the origin, defining a boundary sphere. The surface of this sphere represents all view directions. Thus, let S be the 3D space which represents the bounding sphere of the complete model that contains the set of camera locations or viewpoints. Let $V = \{v_1, v_2, ..., v_N\}$ be the space of $|V|$ camera locations defined as a discrete and finite set of iso-distributed samples over the surface of S. To guarantee a regular distribution we use the HealPix package [GHB*05]. An example of a sphere and the iso-distributed samples over the surface is shown in Fig. 2, highlighting the images related to each viewpoint.
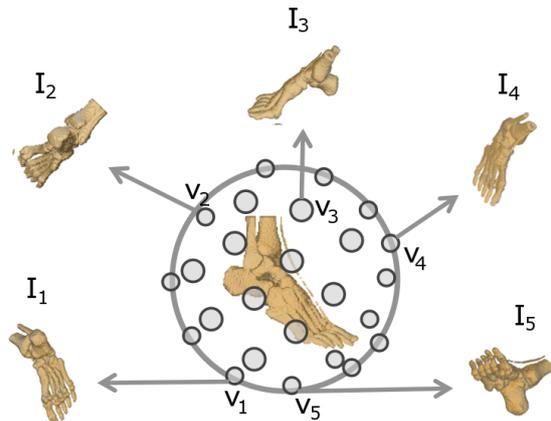
Figure 2: Description of the sampled sphere, S, each image I_i is related to each viewpoint v_i.

We assume that the up vector of the cameras can be arbitrarily chosen at each location due to the camera rolls not having an effect on the visibility. Each view projection of the selected features is a 2D image, I_i. Let $I = \{I_1, I_2, ..., I_N\}$ be the set of all images projected over the sphere and $H(v_i)$ their corresponding entropy. Generally, focused regions can be completely occluded by others, and this fact can difficult the computation of the final entropy of a viewpoint. We can assume that each I_i is the set of layers where each one is the projection of all the voxels of the f_j feature without occluders. These layers could be used to compute the $H_{f_j}(v)$ entropy, and also to define the shape similarity function.

3.2 Shape similarity term

Once we have the camera locations and a reference 2D image I_r projected from the reference camera viewpoint, the objective is to find a complementary image

$I_i \in I$ so that it offers the most complementary information to the reference image I_r. For this task, an image similarity procedure among I_r and each candidate I_i is computed iteratively. In order to perform this estimation, we first need a definition about what we refer for similarity. Although different features and definitions can be considered for this task, we just will use simple image information that can be useful to discriminate similar visual shapes in terms of silhouette and region properties. In particular, we focus on the external contour similarity of 3D object projections as well as the amount of visual information of each volume structure from a particular point of view. As we will see, it allows for a simple estimation that is effective on selecting an useful complementary camera.

The similarity measure used to compute this information is obtained as follows:

$$Sim(I_r, I_i) = (1 - \beta) \cdot Dist(I_r, I_i) + \beta \cdot Corr(I_r, I_i), \quad (2)$$

where *Sim* stands for the similarity measure between a pair of images, *Dist* (see Eq. 3) is the function that computes a shape similarity among a pair of images, and *Corr* (see Eq. 6) is the function that computes the correlation among percentage of visible information for each structure present in both images. Finally, β is a regularization factor. $Sim(I_r, I_i)$ tends to 0 when I_r and I_i are close to each other. On the other hand $Sim(I_r, I_i)$ tends to 1. Note that for simplicity, we consider shape contour and label properties information in order to compute shape similarity.

For the shape similarity term $Dist(I_r, I_i)$, our goal is to obtain a matching cost among two object external contours. This cost is low for similar shapes even if the shape is captured under different rotations or reflections. As shown in Fig. 3, for this task, we first binarize the image. Next, we compute the external contour of the volume by applying a Canny [Can86] edge detector on the binary volume information projected to a particular 2D view. Finally, the contour is vectorized by looking over the contour the distance of each point to the center of mass, see Fig. 3(c). In order to reduce the matching time, we just extract the main external shape silhouette removing inner object contour points from projected 2D shape. We experimentally found that this does not have a negative effect on final performance. Figure 4 depicts the resulting vectorized contour of the foot shown in Fig. 3 where the x-axis corresponds to the position on the contour and the y-axis is the distance of this position to the center of mass. Once the contour is vectorized, it can be efficiently analyzed using dynamic programming. [SK08] presents an indexing approach for time series indexing and mining based on hashing Euclidean distance rules. Although this approach presents promising results for large time series datasets, it is demonstrated that for reduced feature vectors, like the ones we use to codify

the shape silhouette, the estimated warping cost based on dynamic programming uses to obtain more accurate results. In consequence, we use Dynamic Time Warping (DTW). The rationale behind DTW is, given two time series, to stretch or compress them locally in order to make one resemble the other as much as possible. The distance between both time series is computed, after stretching, by summing the distances of individual aligned elements. DTW is an algorithm commonly used for measuring similarity between two sequences which may vary in time or speed. DTW has been applied to video, audio, and graphics -indeed, any data which can be turned into a linear representation can be analyzed with DTW. In this sense, we apply DTW to match the vectorized shapes of each 2D projected volume viewpoint in relation to the reference image.

Reference Image, I_r

(a) Reference 2D image (b) Binary image (c) Contour
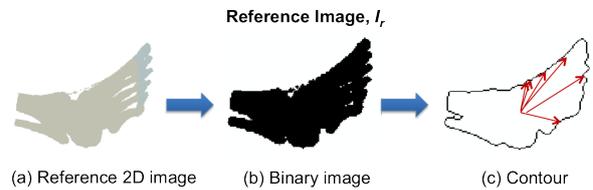
Figure 3: Description of the process for extracting the contour of the reference image.
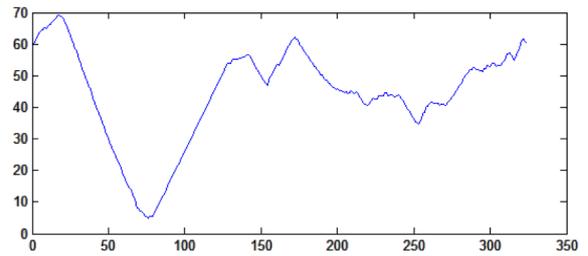
Figure 4: Vectorized contour of the reference image.

In order to obtain reflection invariance we apply the same procedure for the image I_i, and for the horizontal and vertical image reflections, I_i^h and I_i^v, respectively. Figures 5 and 6 depict the process for extracting the contour of a candidate image I_i, and its final vectorized contour, respectively.

Image, I_i, for viewpoint v_i

(a) 2D image (b) Binary image (c) Contour
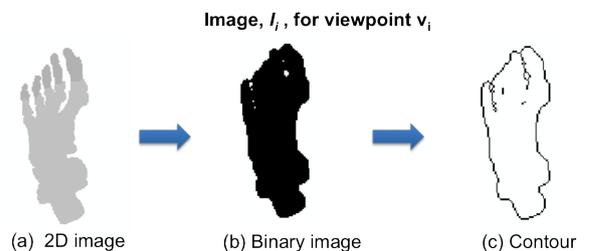
Figure 5: Description of the process for extracting the contour of an image, I_i.

On the other hand, note that rotation and viewpoint invariance will be obtained by the matching procedure
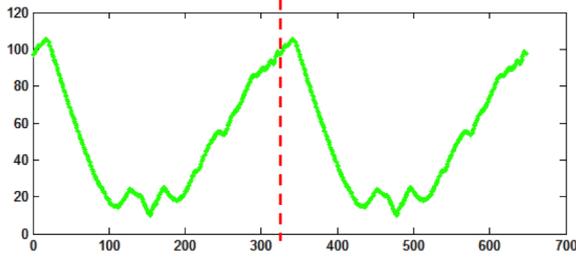
Figure 6: Duplicated vectorized contour of a candidate image I_i to guarantee rotation invariance.

based on Dynamic Time Warping, according to the horizontal an vertical image reflections. Thus, final similarity is obtained as:

$$Dist(I_r, I_i) = \min(d(I_r, I_i), d(I_r, I_i^h), d(I_r, I_i^v)), \quad (3)$$

where d function computes the Dynamic Time Warping alignment between two series. For this task, I_r is decomposed into a one-dimensional sequence, vectorizing the external shape contour of the connected components in I_r as described above, see Fig. 4. The shape vectors of a image, I_i, are duplicated in size, concatenating twice the vectoring shape to guarantee rotation invariance, see Fig. 6. Then $d(I_r, I_i)$ is computed as a Dynamic Time Warping algorithm. Specifically, our DTW algorithm is defined to match distortions between two models, finding an alignment/warping path between the two shape series I_r and I_i. In order to align these two sequences, a $M_{m \times n}$ matrix is designed, where the position (x, y) of the matrix contains the alignment cost (i.e., defined later in Eq. 5) between positions x and y of both series, see Fig.7.
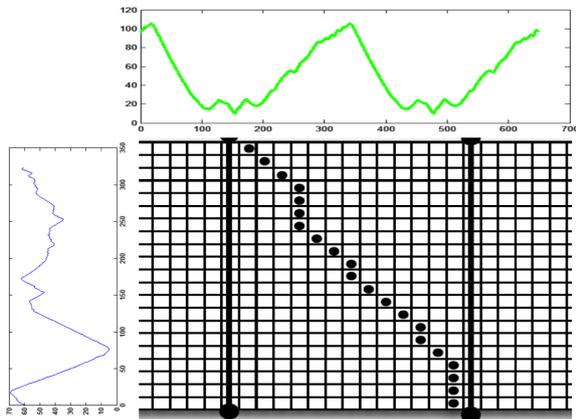
Figure 7: Cost matrix, M, with the cost alignment between positions x and y of both series.

From the matrix, M, a warping path of length T is defined as a set of contiguous matrix elements, defining a mapping between I_r and I_i as $W = \{w_1, ..., w_T\}$, where w_i indexes a position in the cost matrix. This warping path is typically subjected to several constraints:

Boundary conditions: $w_1 = (1, 1)$ and $w_T = (m, n)$.

Continuity and monotonicity: Given $w_{t-1} = (a', b')$, then $w_t = (a, b)$, $a - a' \le 1$ and $b - b' \le 1$, this condition forces the points in W to be monotonically spaced in shape.

We are generally interested in the final warping path that minimizes the warping cost:

$$d(I_r, I_i) = min\left(\frac{M(w_t)}{|W|}\right), \quad (4)$$

where $|W|$ compensates the different lengths of the warping paths. The cost at a certain position $M(x, y)$ can be found as the composition of the Euclidean distance $c(x, y)$ between the feature vectors of the two sequences and the minimum cost of the adjacent elements of the cost matrix up to that point, i.e.,

$$M(x,y) = c(x,y) + \min(M(x-1,y-1), M(x-1,y), M(x,y-1)). \quad (5)$$

The cost $c(x, y) = (x - y)^2$ that is the Euclidean distance between the feature vectors at positions x and y.

In order to detect the beginning and ending positions of the candidate shape that minimizes the matching cost, the current ending cost is checked (the cost of the element in the last row). This minimum value is assigned to $d(I_r, I_i)$, which is used in Eq. 3. Figure 8 depicts in green color the best alignment (i.e., the one that minimizes the matching cost) among the vectorized shapes of the reference image, I_r, and the candidate image, I_i.
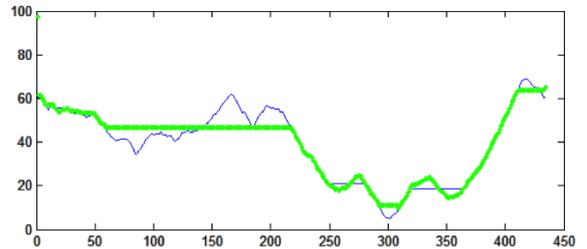
Figure 8: Alignment found using DTW among I_r and the candidate image I_i.

Given that some labeled structures may have similar shape but different resolution, and thus offer different percentage of information, we complement the DTW similarity procedure with a correlation of information from each feature f_j, $f_j \in F$, among two images. The features we use in our foot example are the labels of the volume (i.e, toes, palm and ankle). For this task *Corr* is computed as follows:

$$Corr(I_r, I_i) = \sum_{f_j \in F} \left| \frac{count(I_r, f_j)}{\sum_{f_l \in F} count(I_r, f_l)} - \frac{count(I_i, f_j)}{\sum_{f_l \in F} count(I_i, f_l)} \right|, \quad (6)$$

where $count(I_r, f_j)$ estimates the number of visible pixels of feature f_j in I_r. When transparency is taken into account in the rendering, each projected pixel can contribute several times in the *count* function for different structure labels, as many as non-occluded structures

have been projected onto that pixel. Thus, in Eq. 6 normalization is based on the total amount of visible pixels for each structure. In this sense, similar external shape projections but with different amount of visible information from each structure in relation to the reference image will penalize the similarity measure. On the other hand, similar visual proportion for each structure in comparison with the reference image will be penalized if the external projected structure among views differs. This trade-off regularized by the β term in Eq. 2 allows for a robust estimation of complementary view selection, assigning higher score to those views that offer additional information to the primary view in terms of visual structures and external shape.

3.3 Selecting the optimal complementary camera

Given a reference camera viewpoint $v_r \in S$ and its projected image I_r, we define the function $Dual(v_r)$ as[1],

$$Dual(v_r) = max(Sim(I_r, I_i) \cdot H(v_i)) : \forall I_i \in I. \quad (7)$$

Similarity measure defined in Eq. 2 is used to mask the entropy at each sampled viewpoint location v_r into the sampled space S.
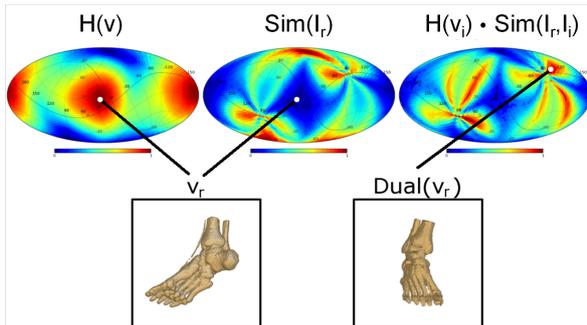
Figure 9: The first row shows the projection of the sampled sphere of each computed viewpoint function. The second row shows the reference image and the complementary camera projection. The entropy $H(v)$, similarity $Sim(I_r, I_i)$ and the final composition $H(v_i) \cdot Sim(I_r, I_i)$ is used to find the complementary viewpoint, $Dual(v_r)$.

Figure 9 describes the final composition to find the optimal complementary viewpoint. The entropy $H(v)$ is computed for all the selected features (or isovalues) (see Eq. 1). Each projected image stores for each pixel the first intersection with the focused feature and the corresponding depth, when this intersection exists. These images are next used to compute the entropy, $H(v)$. Next, given a viewpoint location, v_r, of the sampled sphere, S, the similarity of all the samples related can

be computed using the $Sim(I_r, I_i)$ function. Finally, the maximum value of these former costs locates the complementary camera $Dual(v_r)$.

4 SIMULATIONS

This section describes the design of the experiments, the analysis of the proposal for computing complementary camera viewpoints and the obtained results.

4.1 Design of the experiments

We have analyzed the proposed method using an Intel Core I7 870 with 16 GB memory equipped with an NVidia GeForce GTX 690 GPU with 4GB of GDDR5 memory. The viewport size is 512×512.

In order to test our method, we have sampled the searching space S with 192, 768 and 3072 viewpoint locations. We are able to compute 192 views in 0.03 seconds in average, 768 in 0.15 seconds and 3072 in 0.61 seconds. We have used three data sets: (1) the *Foot*[2] data set is a 128^3-sized CT scan of a human foot; (2) the *Thorax* [3] data set represents a phantom human torso of 400^3 voxels; and (3) the *Walnut*[4] data set, which is a CT scan of a walnut of $400 \times 296 \times 352$ voxels. In the walnut dataset, the structures have a configuration of onion-peel-like, i.e., each structure is totally or partially contained in another structure.

We have selected three structures of the foot (toes, palm and ankle), twelve for the thorax to test different amount of overlapping between structures, and three features for the walnut (i.e., shell, seed, and core segments). The selected structures are based on the public label annotations of the used datasets. We empirically setup the β parameter to 0.5 in order to compute the similarity term.

With the aim of validating the goodness of our proposal, we performed a live-user evaluation. In our trial participated 20 subjects. The users were asked to score a best view and a set of best complementary views in relation to the reference ones. With this experiment we look for the correlation among observers to the best and complementary viewpoint selection based on the mean number of votes per viewpoint. This 3D map serves as a qualitative evaluation in order to look for observers agreement. We also use this information to compare the user selection with our automatic generated 3D score maps for the best and the complementary viewpoints.

4.2 Analysis

In our analysis, the trial performed by the users consisted of the four following steps. Firstly, each trialist rated all the images in the range [1..5]. Secondly,

---

[1] We experimentally found that the product rule obtained better trade off results in comparison to other metrics, such as the sum rule.

[2] http://www.slicer.org/archives.
[3] http://www.voreen.org.
[4] http://www.uni-muenster.de/EIMI/.

among all the images, each trialist selected the one that he considered best reference image (I_r). Next, looking at this reference image and the remaining ones again, they were asked to rate them according to how good the complementary view each one offers is. Finally, they were asked to locate the best complementary image to the one previously chosen as the reference image. For the sake of simplicity, all users were asked to rate the best viewpoint only on the Thorax data set, meanwhile they were asked to rate all the complementary viewpoints for the three data sets. For all the cases of study, 192 images were provided to the participants.

On the other hand, we did similar steps for evaluating our proposal. First of all, for each viewpoint, we rated them using the entropy, $H(v)$. Secondly, we selected as reference image, I_r, the viewpoint with the maximum entropy. Next, we computed the shape similarity term. Finally, we selected the optimal complementary camera for the previously selected reference image, I_r, as described in Section 3.3.

Table 1 depicts the results for the Thorax data set with 192 viewpoint locations for the live-user evaluation and the proposed algorithm. The first observation for the best camera selection on the Thorax data set is that the users tend to rate all those viewpoints that they do not consider relevant very low in the scale, meanwhile they offer high values to those viewpoints they consider relevant. In this sense the densities of the 3D map tend to be a bimodal distribution. In contrast, in our case our automatically computed map tends to be a Gaussian distribution of scores. The important point of this experiment is that there exists a high agreement among observers in relation with the best viewpoint and that this viewpoint (Table 1 (a) and (e)) highly correlates with the one obtained by our automatic method (Table 1 (b) and (f)). Moreover, both distributions of scores of the best view share similar patterns.

In relation to the complementary viewpoints, as depicted in the figures of Table 1 (c) and (d), there also exists a high agreement among observers and a high correlation with our computed scores for complementary viewpoint selection. It is remarkable that for some symmetric volumes, such as the Thorax data set, we still obtain a high score for that view (Table 1 (g) and (h)), although the average complementary view chosen by the users do not correspond to the one selected by our system. This is mainly because we obtain similar scores for opposite viewpoints given the symmetry of the visualized objects, especially when the number of analyzed viewpoints tends to be 3072 views.

In Table 2, we also show the best complementary selection obtained by the live-user evaluation for the Foot and Walnut data sets based on our best primary view selection. Comparing these views with the complementary ones computed by our approach (see Table 3 and 5) one can see a high agreement among the selected views.

Live-User Evaluation   **Proposed Algorithm**

(a)   (b)

(c)   (d)

(e)   (f)

(g)   (h)
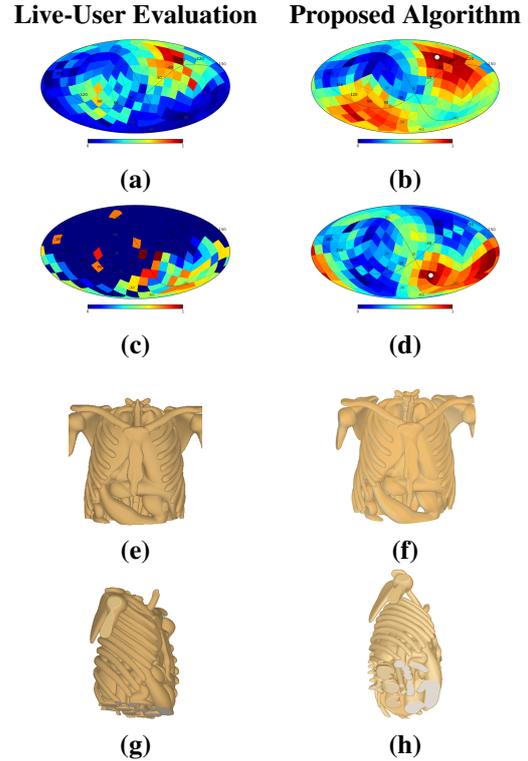
Table 1: Comparison of the results obtained with the live-user evaluation and our proposed method for the Thorax data set with 192 viewpoint locations: (a) and (b) show the initial score distribution, (c) and (d) refer to the complementary view score distribution, (e) and (f) show renderings of the reference image I_r selected by users and computed by our proposed algorithm, respectively, and (g) and (h) show renderings of the complementary view in relation to the same reference image.

(a)   (b)   (c)
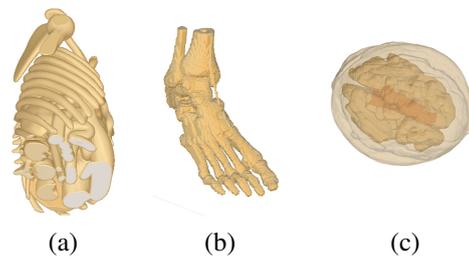
Table 2: Best complementary camera selection by live-user evaluation on the (a) thorax, (b) foot and (c) walnut data sets.

In the tested data sets, the convergence of the complementary camera location directly depends on the variation of the number of samples on the bounding sphere as well as the used searching space.

Tables 3, 4, and 5 show the obtained results for different number of samples in the sphere, $|V|$, from 192 to 3072. The first three rows represent the $H(v)$ entropy of
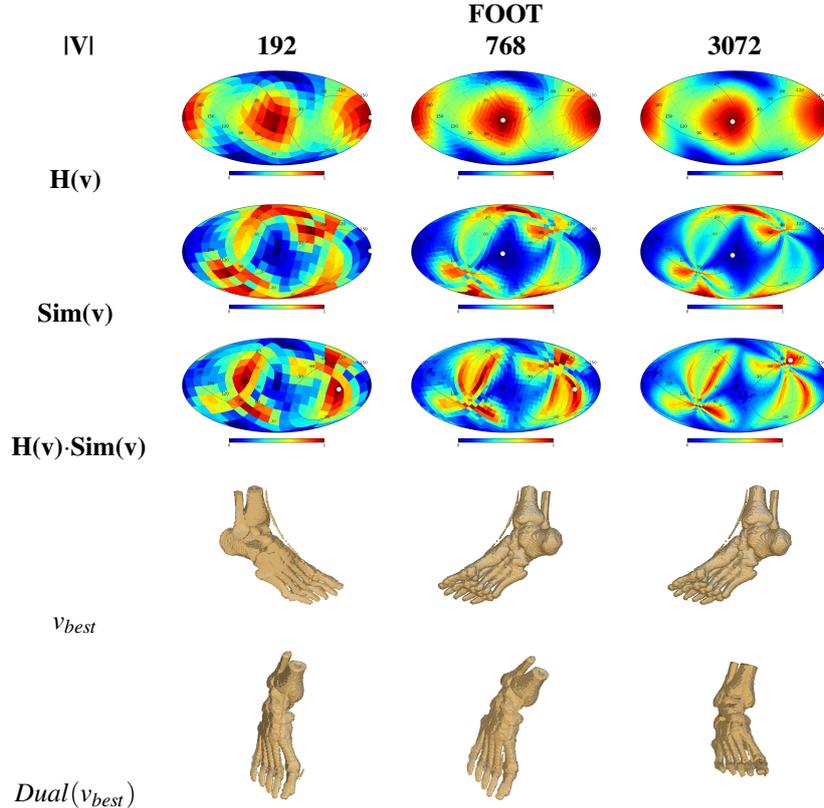
Table 3: A comparison with different number of sampled viewpoints for Foot data set.

the sampled viewpoints over the sphere, the $Sim(v)$ associated the concrete view, and the $H(v) \cdot Sim(v)$ composition, respectively. Each of them is mapped to a plot that represents the whole set of sampled viewpoints in the sphere S. The best viewpoint location in the first row (v_{best}) is highlighted with a white bullet. The similarity plot is computed according to this location ($Sim(v_{best})$) in the second row. Finally, the third row shows the $Dual(v_{best})$ as a white bullet in the composed plot of $H(v) \cdot Sim(v)$. The two last rows are the final renderings associated to v_{best} and $Dual(v_{best})$, respectively.

It can be observed that the computed entropy maps become stable between 192 and 768 cameras. Similarity maps converge from 768 to 3072 sampled positions. Additionally, if we evaluate the final renderings, the 768 is enough in order to obtain a dual camera position with complementary information to the first camera. In this experiment, the reference camera is the best viewpoint located in the entropy map, but we can define any arbitrary position, and then compute the similarity map and its dual function. Note that the complementary camera viewpoints computed highly correlate on average to the ones defined by the live-user evaluation validation.

In addition, we tested the accuracy of the similarity in transparent scenes. Walnut data set is composed of onion-peel-like structures. We analyzed the number of visible pixels of each feature taking into account that several structures can be projected in a pixel. The

count function introduced in Section 3.2 allows one to compute the similarity of this kind of projection. Our method maintains the convergence of the complementary camera between 192 and 768 views for opaque regions as well as the non-transparent structures.

5 CONCLUSIONS AND FUTURE WORK

Traditional research in volume visualization has been focused on involving the user to assist on gathering visual information contained in the data. With this approach users require some expertise, apart from being introduced in a difficult and tedious task in order to setup the parameters of the visualization. In this paper we have described a proposal for obtaining a representative complementary camera in volume rendering without user intervention. Moreover, the proposal is easy to evaluate and robust to changes in the resolution of the final view. Specifically, the proposed method is able to cope with the problem of locating automatically the complementary camera that maximizes the complementary information of the structures. It combines the quantity of information each camera provides for each structure and the shape similarity of the projections of the remaining viewpoints based on Dynamic Time Warping. We have demonstrated the efficacy of our proposal against a benchmark of three data sets. In general, our experiments indicate that the proposal has the po-

V|	THORAX	
	192	768	3072
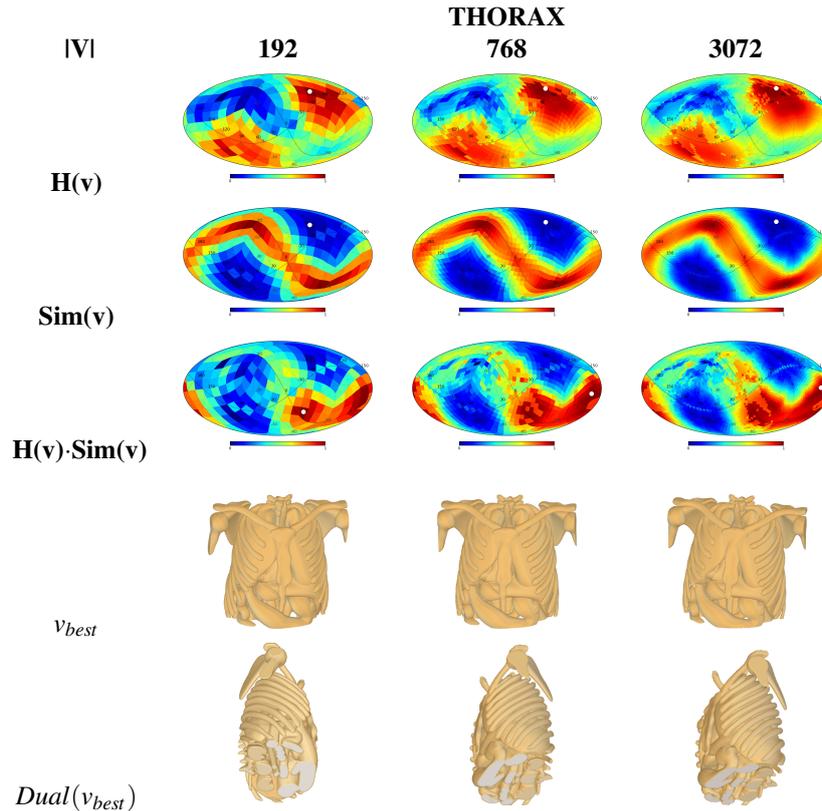
Table 4: A comparison with different number of sampled viewpoints for Thorax data set.

tential to deliver worthwhile similarities in comparison with real users. Our live-user analysis has shown that our approach is close to the preferences of the users. Our results also show that the proposal quickly converges and, with only a 768 sampled positions, it is able to obtain a dual camera position with complementary information to the reference camera. Our future work will focus on exploring new view-dependent descriptors. Currently, we are working on the parallelization of the whole process using OpenCL to get real-time.

ACKNOWLEDGEMENTS

6   REFERENCES

[BG05]   BRUCKNER S., GRÖLLER E.:   Volumeshop: An interactive system for direct volume illustration. In *IEEE Visualization 2005* (Oct. 2005), pp. 671–678.

[BS05a]   BORDOLOI U., SHEN H.:   View selection for volume rendering. In *IEEE Visualization* (2005), p. 62.

[BS05b]   BORDOLOI U., SHEN H.-W.:   View selection for volume rendering. In *IEEE Visualization* (2005), p. 62.

[Can86]   CANNY J.:   A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell. 8*, 6 (June 1986), 679–698.

[CMH08]   CHAN M. Y., MAK W., H.QU: An efficient quality-based camera path planning method for volume exploration. In *Proc. ISVC 2008, LNCS-5359* (2008), et al (eds.) G. B., (Ed.), Springer-Verlag Berlin Heidelberg, pp. 12–21.

[CQWZ06]   CHAN M. Y., QU H., WU Y., ZHOU H.: Viewpoint selection for angiographic volume. In *Proc. ISVC 2006, LNCS-4291* (2006), Springer-Verlag, pp. 528–537.

[GHB*05]   GORSKI K., HIVON E., BANDAY A., WANDELT B., HANSEN F., REINECKE M., BARTELMAN M.: Healpix: a framework for high resolution discretization, and fast analysis of data distributed on the sphere. *The Astrophysical Journal* (2005), 622–759.

[GP09]   GRAU S., PUIG A.:   An adaptive cutaway with volume context preservation. In *Advances in Visual Computing* (2009), vol. 5876 of *LNCS*, pp. 847–856.

[KBKG08]   KOHLMANN P., BRUCKNER S., KANITSAR A., GRÖLLER E.:   The livesync++: Enhancements of an interaction metaphor. In *Graphics Interface*
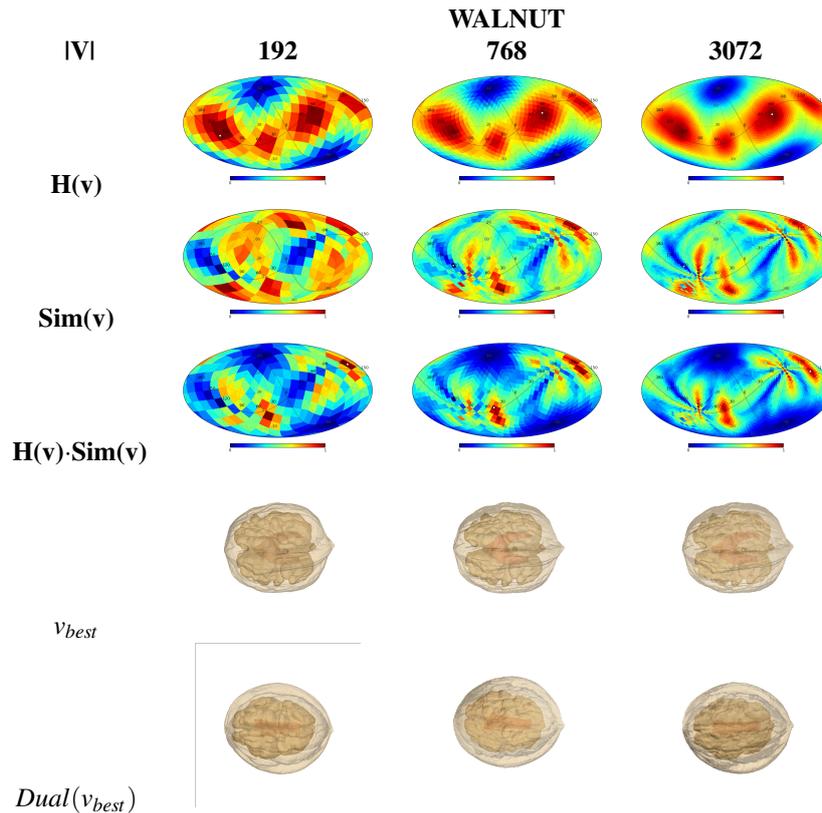
Table 5: A comparison with different number of sampled viewpoints for the Walnut data set.

(08), pp. 81–88.

[KBKG08] KOHLMANN P., BRUCKNER S., KANITSAR A., GRÖLLER M.: The livesync++: Enhancements of an interaction metaphor. In *Graphics Interface* (2008), pp. 81–88.

[MNTP07a] MÜHLER K., NEUGEBAUER M., TIETJEN C., PREIM B.: Viewpoint selection for intervention planning. In *IEEE Symp. on Visualization 2007* (2007), pp. 267–274.

[MNTP07b] MÜHLER K., NEUGEBAUER M., TIETJEN C., PREIM B.: Viewpoint selection for intervention planning. In *Proceedings of Eurographics/IEEE VGTC Symposium on Visualization 2007* (2007), pp. 267–274.

[PPB*05] POLONSKY O., PATANE G., BIASOTTI S., GOTSMAN C., SPAGNUOLO M.: What's in an image? *The Visual Computer 21* (2005), 840–847.

[PPVN08] P. P. VÁZQUEZ E. M., NAVAZO I.: Representative views and paths for volume models. In *Smart Graphics* (2008), pp. 106–117.

[RBGV08] RAUTEK P., BRUCKNER S., GRÖLLER M. E., VIOLA I.: Illustrative visualization: New technology or useless tautology? *SIGGRAPH Comput. Graph. 42*, 3 (2008).

[SK08] SHIEH J., KEOGH E.: isax: Indexing and mining terabyte sized time series, sigkdd. pp, 2008.

[TFTN05] TAKAHASHI S., FUJISHIRO I., TAKESHIMA Y., NISHITA T.: A feature-driven approach to locating optimal viewpoints for volume visualization. *IEEE Visualization 0* (2005), 63.

[VFSG06] VIOLA I., FEIXAS M., SBERT M., GRÖLLER E.: Importance-driven focus of attention. *IEEE Trans. on Visualization and Computer Graphics 12*, 5 (2006), 933–940.

[VKG05] VIOLA I., KANITSAR A., GRÖLLER E.: Importance-driven feature enhancement in volume visualization. *IEEE Trans. on Visualization and CG 11*, 4 (2005), 408–418.

[ZWD12] ZHANG Y.-S., WANG B., DAI C.-J.: Viewpoint Selection Based on NM-PSO for Volume Rendering. In *Intel. Comp. Theories and Applications*, vol. 7390. 2012, pp. 487–494.