Blood Flow Clustering and Applications in Virtual Stenting of Intracranial Aneurysms

Steffen Oeltze, Dirk J. Lehmann, Alexander Kuhn, Gábor Janiga, Holger Theisel, and Bernhard Preim

Abstract—Understanding the hemodynamics of blood flow in vascular pathologies such as intracranial aneurysms is essential for both their diagnosis and treatment. Computational fluid dynamics (CFD) simulations of blood flow based on patient-individual data are performed to better understand aneurysm initiation and progression and more recently, for predicting treatment success. In *virtual stenting*, a flow-diverting mesh tube (*stent*) is modeled inside the reconstructed vasculature and integrated in the simulation. We focus on steady-state simulation and the resulting complex multiparameter data. The blood flow pattern captured therein is assumed to be related to the success of stenting. It is often visualized by a dense and cluttered set of streamlines. We present a fully automatic approach for reducing visual clutter and exposing characteristic flow structures by clustering streamlines and computing cluster representatives. While individual clustering techniques have been applied before to streamlines in 3D flow fields, we contribute a general quantitative and a domain-specific qualitative evaluation of three state-of-the-art techniques. We show that clustering based on streamline geometry as well as on domain-specific streamline attributes contributes to comparing and evaluating different virtual stenting strategies. With our work, we aim at supporting CFD engineers and interventional neuroradiologists.

Index Terms—Blood flow, aneurysm, virtual stenting, clustering, evaluation

1 Introduction

Intracranial aneurysms, also referred to as cerebral aneurysms, represent a pathological, balloon like dilation of cerebral vasculature due to a weakness of the arterial wall. They occur with a prevalence of about 2 percent in Western Europe [1]. Their rupture is associated with a mortality rate of ≈ 50 percent. Among other treatment options, *stenting* plays an increasingly important role. In stenting, the flow is diverted around the aneurysm by an expandable mesh tube (*stent*), thereby reducing and decelerating its inflow (Fig. 1a). The blood flow pattern is among the hemodynamical parameters that are assumed to be related to the success of stenting [2], [3], the development of thrombosis (blood clotting, which is a desirable outcome of stenting) [4], and the risk of aneurysm rupture [5]. A better understanding of these relations may contribute to patient selection for flow diverting stents. While they often lead to thrombosis and reverse remodeling, adverse effects leading to late rupture were also observed [3]. With the increased number of treatment options and available types of stents, the need for decision support is strongly increased.

Computational fluid dynamics (CFD) simulations, which generate patient-specific hemodynamic data, are employed to better understand the effect of stents on aneurysmal hemodynamics and for predicting treatment success [2], [6], [7]. In *virtual stenting* (VS), different types of stents are modeled at different locations inside the reconstructed vascular anatomy and integrated in the simulation. We focus on steady-state simulations since major aspects of aneurysmal hemodynamics may be inferred from steady flow [8]. The simulation results in a complex multiparameter data set comprising several scalar and vectorial attributes. The blood flow pattern captured therein, is often visualized for investigation by a dense and cluttered set of streamlines colored according to one of the scalar attributes.

We present a fully automatic approach for reducing visual clutter and exposing characteristic flow structures by grouping similar streamlines and computing group representatives. We quantitatively evaluate three conceptually different techniques for the grouping: *k-means* clustering, *agglomerative hierarchical clustering* (AHC) in four variations (single link, complete link, average link, and Ward's method), and *spectral clustering* (SC). While each individual technique has been applied to streamlines in 3D flow fields [9], [10], [11], [12], the quality of their results has not been compared before. The gained insight is valuable for all applications employing streamline clustering.

Cluster representatives, which summarize the complex blood flow, are derived from the clustering result. We adapt a type of representative that is employed in clustering fiber tracts of the human brain. In a qualitative expert evaluation of visual blood flow summaries, we compare the quantitatively best performing clustering techniques and the corresponding representatives. Furthermore, we show that clustering streamlines also based on domain-specific attributes supports the evaluation of virtual stenting strategies. For instance, clustering based on the local

- *S. Oeltze, D.J. Lehmann, H. Theisel, and B. Preim are with the Department of Simulation and Graphics, University of Magdeburg, Magdeburg, Germany. E-mail: {stoeltze, dirk, theisel, bernhard}@isg.cs.uni-magdeburg.de.*
- *A. Kuhn is with the Zuse Insitute, Berlin, Germany. E-mail: kuhn@zib.de.*
- *G. Janiga is with the Institute of Fluid Dynamics and Thermodynamics, University of Magdeburg, Magdeburg, Germany. E-mail: janiga@ovgu.de.*

Manuscript received 26 Dec. 2012; revised 10 Oct. 2013; accepted 17 Dec. 2013; date of publication 8 Jan. 2014; date of current version 5 Mar. 2014. Recommended for acceptance by R. Machiraju. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TVCG.2013.2297914

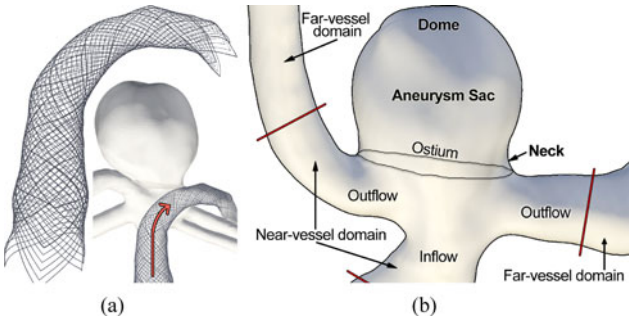

Fig. 1. (a) Flow diverting stent and its deployment (arrow indicates flow direction). (b) Morphological features of a saccular aneurysm (bold) and subdivision of the surrounding vascular domain (red lines).

residence time (RT) of blood flow within the aneurysm gives hints on potential locations of thrombosis initiation.

In summary, our contributions are:

- Quantitative evaluation of three conceptually different streamline clustering techniques
- Visual summary of flow patterns and design lessons
- Expert evaluation of visual flow summaries
- Application-specific insight from clustering domain-specific streamline attributes
- A tailor-made type of cluster representative

We aim at supporting CFD engineers in investigating simulation results. In a dense sampling of aneurysmal flow by thousands of streamlines, they rely on filtering these lines, a locally restricted streamline seeding or on global hemodynamic parameters. Minor, local changes of the flow pattern yet influencing the success of stenting, may remain unnoticed. We further aim at supporting interventional neuroradiologists in developing a patient-specific treatment strategy. CFD results are not yet part of the clinical routine. Hence, the physicians have little experience in investigating flow data. Our visual flow summary simplifies the access to flow data, it is easy to read, and it contributes to the communication between CFD engineers and physicians. We employ our approach amongst others to data of the virtual intracranial stenting challenges (VISC) in 2009 and 2010.

2 Medical and Technical Background

This section briefly overviews the treatment of intracranial aneurysms, introduces the research field virtual stenting, and describes our data generation pipeline.

2.1 Treatment of Intracranial Aneurysms

Intracranial aneurysms usually develop somewhere at the *Circle of Willis*. Their shape may be characterized as saccular, fusiform or dissecting with saccular having by far the highest prevalence [13]. The morphological features of a saccular aneurysm are illustrated by Fig. 1b. Most aneurysms remain undetected until rupture. While surgical clipping has been the gold standard in treatment for decades, the number of endovascular interventions is increasing. They bear less intraoperative risk and may be applied, e.g., by an interventional neuroradiologist, to aneurysms which are difficult or impossible to reach for a surgeon [14]. In *coiling*, the aneurysm is filled with platinum coils to promote thrombosis, which may eventually seal the aneurysm. Self expanding, high-profile, flow-diverting stents provide a promising alternative to coiling in patients with complex aneurysms (Fig. 1a). They reduce and decelerate the blood circulation into the aneurysm, thereby causing a prolonged residence time, which in turn promotes thrombosis formation [13].

Despite the progress in interventional techniques, the associated risks persist, e.g., injury of the aneurysmal wall during stent insertion. A detailed risk and benefit estimation and a deeper insight into the hemodynamics of blood flow that cause aneurysm development and rupture are necessary.

2.2 Virtual Stenting

Virtual stenting is a collaborative effort between CFD engineers, physicians, and computer scientists. Its main objectives are supporting clinical decision making and stent design. In the former, questions such as "Is the vascular and aneurysmal morphology eligible for stenting?" and "Which stent should be used and where should it be placed?" need to be answered. In stent design, different properties, e.g., grade of mesh porosity and strut size, and their impact on the hemodynamics of blood flow are investigated.

One challenge in VS is comparing results of different CFD simulations, e.g., before and after stenting [6]. We support a comparison by visual summaries of blood flow. So far, it is often based on global values such as aneurysmal inflow rate [15]. Sometimes, the aneurysm wall is colored according to a hemodynamic parameter and presented in a side-by-side view [6]. Streamlines are employed for comparing flow patterns. They are often seeded on the ostium and displayed side-by-side [2], [6], [16]. However, either the entire set of lines is displayed leading to visual clutter or representative lines must be selected manually.

2.3 Hemodynamic Data Generation Pipeline

We briefly summarize our hemodynamic data generation pipeline (see [15], [17] for details). First, image data of the aneurysm morphology including the vasculature in the close surrounding are acquired, e.g., by 3D rotational angiography or computed tomography (CT) angiography. Next, the aneurysm and the vasculature are segmented via thresholding. Afterwards, a surface mesh of the vessel wall is reconstructed from the segmentation result and optimized [18]. Then, the *ostium* is extracted [19]. It separates the aneurysm from the parent vessel and approximates the original vessel wall (Fig. 1b). It is frequently used to explore the flow into the aneurysm, e.g., by seeding streamlines there [20]. Next, the stent geometry is modeled and deployed to the vessel wall. Finally, a volume mesh is constructed based on the surface meshes of the vessel wall and the stent using ANSYS IcemCFD (Ansys Inc., Canonsburg, PA). Fluid flow simulations are performed in ANSYS Fluent 12 (Ansys Inc., Canonsburg, PA).

3 Related Work on Partition-Based Flow Visualization

Flow visualization techniques have been categorized by Post et al. [21] into direct, texture-based, geometric, and

feature-based techniques. Salzbrunn et al. [22] added the class of partition-based techniques, which decompose a flow field based on vector values, integral curve properties or topological features. Blood flow clustering based on vector values has been presented in the context of cardiac blood flow [23]. However, we follow the arguments in [10] and advocate the use of integral curves since they represent continuous flow patterns traced over the domain instead of a very local vectorial flow information. We briefly recapitulate approaches for flow decomposition based on integral curves and classify them into user-guided and automatic partitioning. For State-of-the-Art reports on topology-based decomposition and visualization of flow, see [24], [25].

3.1 User-Guided Flow Partitioning

The approaches in this class decompose a set of integral curves guided by the user. Salzbrunn and Scheuermann [26] propose combined Boolean predicates based on predefined scalar quantities, which determine for each streamline whether it has a desired property. Predicates on pathlines are applied to the visual analysis of measured blood flow in aortic aneurysms [27]. A residence time predicate is used for evaluating blood clotting. In [28], a visual analytics approach is proposed for filtering pathlines based on local and global pathline attributes, e. g., curvature and Lyapunov exponent. Pobitzer et al. [29] demonstrate the application of dimension reduction to the set of attributes in order to detect relevant, independent ones. Two other approaches let the user specify interesting integral curves or curve parts in observation instead of attribute space. Advanced virtual probing of measured cardiovascular flow by seeding integral curves on a flexible probing geometry is presented in [30]. Gasteiger et al. employ a lens metaphor for generating focus-and-context visualizations of streamline parts [17].

The lens metaphor facilitates only a local and view-dependent inspection of the flow pattern. It emphasizes or attenuates all streamline parts inside the lens but it does not reduce visual clutter with respect to the flow pattern. Neither lens nor virtual probing deliver reproducible and quantifiable results. Line predicates and the visual analytics of pathline attributes require the user to define attributes and attribute value ranges of interest in order to compose sets of lines, which are homogenous with respect to a certain attribute or a combination of attributes. Automatic flow partitioning approaches employ a data-driven strategy for creating such sets and are hence self-tuning with respect to differences in the flow across aneurysms.

3.2 Automatic Flow Partitioning

Our work is strongly related to approaches, which automatically partition a set of integral curves by means of *clustering*, i.e., grouping similar curves. These approaches differ in the clustering technique and in the similarity measure. Chen et al. propose a two-stage k-means clustering [9]. The initial rough geometry-based partitioning is refined by taking vector and shape properties into account. Both stages are based on euclidean distance as the similarity measure. Cluster representatives are the streamlines closest to the cluster centroids. In [12], agglomerative hierarchical clustering with average link has been used for partitioning. The authors propose a similarity measure that facilitates an interactive, cluster-based exploration of flow with seeding rakes. A saliency-guided streamline seeding is followed by AHC with single link in [10]. Streamlines at cluster boundaries are displayed as representatives. Gasteiger et al. employ local streamline properties to identify and group lines that constitute the *inflow jet*, which is correlated with aneurysm rupture [31]. Rössl and Theisel discuss a *spectral embedding* of streamlines [11]. They demonstrate spectral clustering in the embedding space and compare various similarity measures. Similar to the clustering of integral curves is the clustering of fiber tracts extracted from diffusion tensor imaging (DTI) data. In [32], fiber tracts are partitioned by means of a specialized SC approach. Three types of cluster representatives are investigated in [33]. Moberts et al. evaluate three variants of AHC and four similarity measures for clustering fiber tracts [34]. A new similarity measure in conjunction with AHC using single link is introduced in [35].

AHC, k-means, and SC are the most widely used techniques for clustering streamlines (and fiber tracts). However, the quality of their results in this context has not been individually assessed and compared. We quantitatively evaluate the three techniques, including four AHC variants, by means of internal cluster validity indices (Section 5.4). In a qualitative expert evaluation of the best performing techniques, we identify the most appropriate one for clustering blood flow (Section 6.4). While the clustering in related work is mostly restricted to streamline geometry and derived geometrical attributes, we extend it to domain-specific attributes. We adopt the idea of cluster representatives for reducing visual clutter and assess the approaches in [33].

4 Streamline Generation & Similarity

In this section, we describe our generation of streamlines, their properties, and our streamline similarity measures.

4.1 Domain, Tracing, and Line Properties

The input of the streamline generation is the volume mesh from the CFD simulation (Section 2.3). It is represented as an unstructured grid composed of tetrahedral cells. A vector is stored at each cell point. Before streamlines are generated, the mesh is manually cropped such that it contains only the aneurysm and the *near-vessel domain* [20] (Fig. 1b). This enables us to focus the analysis and strongly improves the expressiveness of the clustering. It is very likely that streamlines follow a similar course in the feeding vessel (inflow) and they may also follow a similar course in a draining vessel (outflow). However, depending on where they enter the aneurysm, their course may strongly differ inside. If the *far-vessel domain* was also considered in clustering, these differences would have less impact.

To assess the in- and outflow of the aneurysm, streamlines have been seeded on the ostium. The ostium is represented by a triangle mesh whose vertices have been homogeneously distributed such that the under- and

overrepresentation of flow parts are avoided [19]. The number of vertices is adjusted such that the mesh resembles the former vessel wall. Streamlines were traced in ParaView (Kitware, Clifton Park, NY). A fifth order Runge-Kutta method has been employed with an integration step size that was constantly adjusted according to an estimated error. The tracing was carried out in backward and forward direction. The resulting two lines were merged such that a linear traversal of the vertices from in- to outflow is possible.

Line Properties: The streamlines differ in their number of vertices and in their length. The former has a strong impact on the computational time of most inter-streamline similarity measures. The similarity itself is strongly influenced by streamline length. Two lines may follow a similar course for a long time but then, one of them is terminated. Most similarity measures assign a much higher weight to the difference in length than to the similarity over a long run. In all our data sets, a few lines follow a course very similar to a large set of neighboring lines but are considerably shorter. They occur close to the vessel wall due to early termination of the integration. We consider them as incomplete rather than incorrect data entities. Hence, the clustering should group them with the streamlines having a similar course. Still, we term them outliers in the following.

4.2 Geometry-Based Streamline Similarity

Geometry-based streamline similarity (or dissimilarity) is often expressed by a distance measure. The choice of a measure depends on the application. General requirements are positive-definiteness and symmetry. A valid example is the Hausdorff distance. However, this distance is very sensitive to streamline length, since it outputs the maximum of pointwise distances [11]. A less sensitive measure is the *mean of closest point distances* (MCPD) [36]:

$$\begin{aligned} d_M(s_i, s_j) &= mean(d_m(s_i, s_j), d_m(s_j, s_i)) \\ with\ d_m(s_i, s_j) &= mean_{p_l \in s_i} \min_{p_k \in s_j} \|p_k - p_l\|. \end{aligned} \quad (1)$$

Moberts et al. evaluate four similarity measures for clustering fiber tracts and favor MCPD [34]. Yu et al. apply MCPD for clustering streamlines and report that the clusters comprise important flow features [10]. In [11], five similarity measures adopted from the clustering of fiber tracts are evaluated for clustering streamlines. The rather qualitative evaluation includes MCPD and shows no drawbacks compared to the other measures. In [12], a new similarity measure is compared to three other measures including MCPD. The new measure performs one to two orders of magnitude faster but no advantage in terms of cluster quality is reported. However, MCPD is subjectively rated as producing good quality clusterings. We adopted MCPD and applied it to blood flow clustering. Initial tests showed good results but also revealed that MCPD is still too sensitive to streamline length, in particular when being used with clustering techniques being sensitive to outliers (Table 1). Very small-sized, outlier-corrupted clusters were generated whose representatives distorted the flow summary. We further reduce MCPD's sensitivity by replacing the outer mean in Equation (1) by a minimum computation:

$$\begin{aligned} d_M(s_i, s_j) &= min(d_m(s_i, s_j), d_m(s_j, s_i)) \\ with\ d_m(s_i, s_j) &= mean_{p_l \in s_i} \min_{p_k \in s_j} \|p_k - p_l\|. \end{aligned} \quad (2)$$

If two lines are very similar but one is shorter, d_m from the shorter to the longer line is chosen. The resulting high similarity increases the chance of being assembled.

4.3 Attribute-Based Streamline Similarity

Besides streamline geometry, we employ streamline attributes for clustering. They describe **(1)** the underlying vector field, **(2)** line bending or **(3)** domain-specific aspects:

1. pressure, velocity magnitude, velocity gradient magnitude, angular velocity, vorticity magnitude
2. curvature, torsion
3. distance to ostium, distance to aneurysm wall, local residence time

In the following, we focus on the domain-specific attributes **(3)** since their clustering revealed the most interesting aspects. The distance to the ostium is computed in order to separate flow structures that occur close to the aneurysm's neck from those that occur close to its dome (Fig. 1b). The distance to the aneurysm wall is determined in order to separate flow close to the wall from flow close to the center. Both are inspired by discussions with a neuroradiologist and by clinical research results such as a close correspondence between near-wall flow and wall-shear stress (WSS). They have been computed only at streamline vertices located inside the aneurysm as the distance between the vertex and its closest point (not vertex) on the respective surface.

The residence time of flow inside the aneurysm is crucial in thrombosis formation [4]. We compute it by aggregating partial timing results along each streamline. For each line segment inside the aneurysm, the two associated velocity magnitudes are retrieved from the data. Based on their difference and the segment length, the partial residence time is computed. If a line segment intersects the ostium, the velocity is interpolated at the intersection point. While the other streamline attributes are computed per vertex, the residence time is a single scalar per line.

What is left is the definition of a streamline similarity measure on the attributes. For the local residence time, we employ the absolute difference of two scalars. For the remaining attributes, we first compute a simple statistic that approximates the attribute information along a streamline, e.g., minimum, maximum, mean, or median. Since this breaks down the information to a scalar value, we can apply the same similarity measure as for the residence time.

5 Streamline Clustering Techniques: A Quantitative Evaluation

This section is dedicated to the quantitative evaluation of techniques often used for clustering streamlines (Section 3.2). It starts with descriptions of agglomerative hierarchical clustering and k-means based on [37] and an introduction to spectral clustering based on [38].

TABLE 1
Comparison of Clustering Algorithms with Respect to the Type of Objective Function and the Capabilities to Handle Arbitrarily-Shaped Clusters, Clusters of Significantly Different Size, and Outliers

Property	SC	Agglomerative Hierarchical Clustering				k-means
		Single	Compl.	Avg.	Ward	
OF	global	local	local	local	local	global
Shape	+	+	−	−	−	−
Size	*o*	+	−	*o*	*o*	−
Outlier	+	−	*o*	*o*	+	−

5.1 Agglomerative Hierarchical Clustering

AHC starts with each streamline being a cluster and then, repeatedly merges the two closest clusters until a single cluster is formed. The resulting hierarchy is stored and may be visualized by a *dendrogram*. All merge steps rely on a squared, symmetric distance matrix $\mathbf{M}$ and a measure of cluster proximity. In our case, $\mathbf{M}$ contains the pairwise inter-streamline distances (Eq. (2)). Various cluster proximity measures have been published among which *single link*, *complete link*, *average link*, and *Ward's method* are most popular. In single link, the proximity of two clusters is defined as the minimum distance between any two points in the different clusters. This approach can handle clusters of arbitrary shape, it tolerates considerable differences in cluster size but it is sensitive to outliers. Furthermore, it is infamous for the *chaining effect* leading to clusters containing very dissimilar elements which are connected by a chain of similar elements via some transitive relationship. In complete link, the proximity of two clusters is computed as the maximum distance between any two points in the different clusters. Complete link is less susceptible to outliers but tends to break large clusters and it favors globular cluster shapes. Average link is an intermediate approach between single and complete link. It also strives for globular compact clusters [39]. The proximity of two clusters is defined as the average proximity between pairs of points in the different clusters. Ward's method aims at minimizing the total within-cluster variance. It defines the proximity of two clusters as the sum of squared distances between any two points in the different clusters (SSE: sum of the squared error). Due to the SSE-based proximity, Ward's method favors globular clusters. It was shown to prefer clusters with similar size and to be robust against outliers in the context of 2D curves [40].

All AHC variants lack a global objective function (OF) to be optimized (Table 1). They decide locally which clusters are merged. These decisions cannot be undone such that bad decisions, e.g., involving outliers, are propagated throughout the entire clustering process. A strength of AHC is its ability to rapidly generate different numbers of clusters k by cutting the cluster hierarchy at respective levels. Furthermore, it is non-parametric except for k and the proximity measure. Both strengths explain its frequent use when the "correct" number of clusters is unknown. The user then sequentially browses through the levels. Visually comparing consecutive clustering results is simplified by the locally restricted change (split/merge). AHC's bottleneck in terms of time complexity is the computation of $\mathbf{M}$, which often requires a vast number of euclidean distance tests.

5.2 k-Means Clustering

k-means requires an a priori definition of the number of clusters k by the user. Then, k initial cluster centroids are chosen, often by a random selection of k data entities. Each entity is now assigned to the closest centroid, e.g, by comparing squared euclidean distances. Finally, each centroid is updated to the mean of its assigned data entities (which rarely corresponds to an existing entity). The assignments and updates are repeated until the goal of a global objective function has been achieved. For squared euclidean distances, the objective function usually aims at minimizing the sum of the squared distances of data entities to their cluster centroid (SSE).

Streamlines cannot be directly plugged into k-means since the computation of their mean is undefined. Feature vectors must be derived representing the lines in a new n-dimensional space. A straightforward approach is to use the 3D coordinates of their vertices. Since the number of vertices varies (Section 4.1), each line must be equidistantly resampled to a uniform number. We employ the average number of vertices of all streamlines. A lower-dimensional alternative has been proposed by Chen et al. [9]. Two scalar streamline entropy measures together with the coordinates of start-, middle, and endpoint of the line constitute an 11-dimensional feature vector. Contrary to [9], we employ all dimensions in a single clustering stage since the proposed two stages hamper a user-defined choice of k. However, the latter is required for our quantitative evaluation.

k-means is often computationally faster than AHC since it does not require the computation of pairwise distances between data entities. However, it is sensitive to outliers and fails in handling non-globular clusters and clusters of widely different sizes (Table 1). Its results are dependent on the random initialization of the centroids. A "bad" choice causes the algorithm to get stuck in a local minimum of the objective function. We mitigate this problem by running the algorithm ten times and choosing the result with the minimum SSE.

5.3 Spectral Clustering

Spectral Clustering maps the original streamlines to a *spectral embedding* space where each line is represented by a point (Fig. 2). Key features of the mapping are the preservation of local distance relations between nearby lines and the enhancement of the data's cluster properties, i.e., an

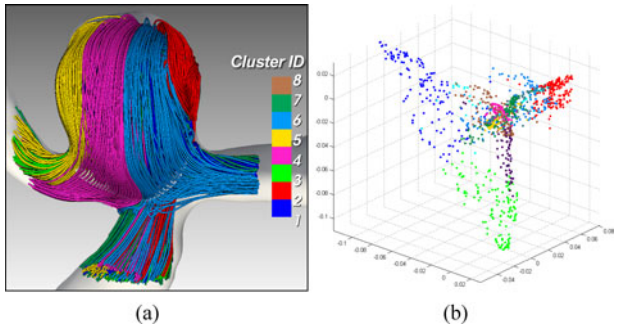

Fig. 2. (a) Spectral Clustering of streamlines in a basilar tip aneurysm. (b) Spectral embedding of the lines. The first three largest eigenvectors are shown.

improved cluster separability. In the following, we use the terms distance and difference interchangeably.

SC can be formulated as a graph partitioning problem [41]. Streamlines are represented by a weighted, fully-connected, undirected graph. The nodes are the streamlines and the edge weights are computed according to Equation (2). The weights are then transformed from difference to affinity such that similar streamlines have a high and dissimilar a low pairwise affinity. Next, the graph is partitioned into subgraphs. Shi and Malik [41] propose to use a *normalized cut* which minimizes the sum of weights of the edges that need to be removed (cut) and at the same time balances the sum of edge weights of the partitions. While this problem is NP hard, a relaxed version is solved by spectral graph partitioning using *Graph Laplacians*.

Given a data set S with n streamlines as a graph and a number of clusters k, **(1)** the $n \times n$ distance matrix $\mathbf{M}$ is computed by a pairwise application of Equation (2) to the lines in S. The same matrix is employed for AHC (Section 5.1). **(2)** Based on $\mathbf{M}$, the $n \times n$ weighted adjacency matrix of the graph is constructed by applying a function f to the entries of $\mathbf{M}$ that gives high values in case of small differences and converges to zero for high differences. The resulting matrix $\mathbf{W}$ is referred to as *affinity matrix*. As f, the Gaussian similarity function is used:

$$f(m_{ij}) = f(m_{ji}) = \exp(-(m_{ij})^2/(2\sigma^2)). \tag{3}$$

The parameter σ controls the width of f thereby steering how rapidly the affinity falls off. **(3)** Next, a $n \times n$ diagonal degree matrix $\mathbf{D}$ is constructed with each diagonal entry d_{ii} being the degree of the node that represents streamline i in the graph. The degree is computed as the sum of weights of the edges incident to the node. **(4)** Now, the *normalized Graph Laplacian* $\mathbf{L}$ is computed [41]: $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{W}$ with $\mathbf{I}$ being the identity matrix. **(5)** Then, the eigenvectors and eigenvalues of $\mathbf{L}$ are determined. The eigenvectors corresponding to the smallest k eigenvalues are used for clustering. **(6)** Let $\mathbf{U}$ be the $n \times k$ matrix that contains the k eigenvectors as columns. Each row i of $\mathbf{U}$ then represents the coordinates of a point that corresponds to streamline i in the $\mathbb{R}^k$ spectral embedding space spanned by the eigenvectors. **(7)** In the embedding, clusters can be detected, e.g., by k-means or an eigenvector rotation [42]. We employ the latter since it suggests an optimum number of clusters based on a user-defined range for k. Since it is based on the largest eigenvectors of $\mathbf{L}$, we change the formulation of $\mathbf{L}$ to:

$$\mathbf{L} = \mathbf{D}^{-1}\mathbf{W}. \tag{4}$$

Local scaling: Zelnik-Manor and Perona propose a local determination of σ since a global value (Eq. (3)) only works well if all clusters are of the same density [42]. Since we cannot guarantee this for our streamlines, we adopt their *local scaling*. A local σ_i is computed for each line i based on the difference between i and its N'th neighbor. A value of $N = 7$ is reported to give good results [42]. However, our experiments indicated that N must be adjusted to each data set. In very dense sets of streamlines, SC partially failed to separate clusters. With increasing density, the local neighborhood of a line contains an increasing number of very similar lines. However, the number of neighbors with an affinity $\gg 0$ should not be "too small and not too large" for SC to work properly [38]. Based on 10 data sets, we identified $N = 5$ percent of the streamline count as appropriate.

SC strives for a globally optimal partitioning while AHC is bound to locally optimal decisions (Table 1). It can handle arbitrary cluster shapes while most AHC variations and k-means favor globular shapes. SC with local scaling considers the local streamline density. This is useful, e.g., if streamlines are seeded with a higher density close to the aneurysm wall. Our implementation of SC is parameter-free except for the range of values for k. Since the eigenvector rotation computes all partitionings within this range, the user can browse also the suboptimal results. SC is biased towards clusters of similar size due to the balancing of edge weights in the graph cutting. On the other hand, this property makes it robust against outliers which was acknowledged in the context of fiber tract length [35]. As for AHC, the bottleneck of SC is the computation of $\mathbf{M}$.

5.4 Quantitative Evaluation

We quantitatively evaluated four variants of agglomerative hierarchical clustering, k-means, and spectral clustering for clustering streamlines. The evaluation was based on five clinical cases together comprising 10 data sets and representing the prevailing types of aneurysms: basilar tip and side-wall aneurysms. Three cases were simulated without virtual stenting (two are shown in Figs. 2a and 4a). Two cases have been simulated before and after stenting, one of them with two types of stents in two different positions (Sections 7.1 and 7.2). The streamline count was between 1,138 and 2,929. The evaluation was restricted to geometry-based clustering (Section 4.2). For each combination of clustering algorithm ($n = 6$) and data set ($n = 10$), streamlines were clustered with the number of clusters being in the range $[2, 20]$ ($n = 19$). This resulted in $6 \times 10 \times 19 = 1,140$ partitionings.

Different measures for assessing the quality of a clustering result have been proposed. In the absence of a ground truth, e.g., external labels provided by an expert, unsupervised measures of cluster validity are appropriate [37]. They are also called *internal validity measures* since they are purely based on information present in the data. We employ four internal measures which together cover the most important aspects of cluster quality [39]:

- *Silhouette Width*: Non-linear combination measure of cluster cohesion and separation. Values are in the range $[-1, +1]$ and should be maximized.
- *Connectivity*: Local measure reflecting to which degree the L most similar neighbors of a streamline are placed in the same cluster. Values are in the range $[0, +\infty]$ and should be minimized. We define $L = 20$.
- *Hubert's* Γ *Statistic*: Measure of correlation between the distance matrix $\mathbf{M}$ and an idealized distance matrix (distance is 0 for streamlines in the same cluster and 1, otherwise). Values of the normalized statistic are in the range $[-1, +1]$ and should be maximized.
- *Stability*: Measure reflecting the stability and hence, the significance of the clusters. Random overlapping

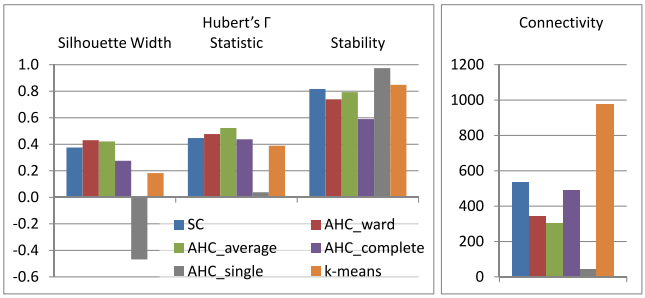

Fig. 3. Average internal cluster validity measures based on 10 data sets. Spectral Clustering, four variants of agglomerative hierarchical clustering, and k-means are compared.

subsamples of the data are repeatedly drawn and clustered using the same algorithm. We draw 20 subsamples. Their clusters are then compared to the partitioning of the original data via the Adjusted Rand Index whose values are in the range $[-1, +1]$ and should be maximized [43].

To ensure comparability of the algorithms, all measures were computed in 3D streamline space although k-means and SC cluster in different spaces, i.e., in feature vector space and in the spectral embedding. The first three measures employ the similarity of two streamlines which is inferred from the distance matrix $\mathbf{M}$. Clustering by k-means has been based on two types of feature vectors (Section 5.2). The type based on streamline resampling consistently achieved better internal measures, which is likely due to the very sparse representation of the streamline course by the other type (only three vertices). Hence, we restrict the presentation of evaluation results to the former. For each algorithm, the internal validity measures were averaged over the 19 partitionings and the 10 data sets (Fig. 3).

Silhouette Width: AHC with single link exhibits a very poor silhouette width (-0.47). This is due to the chaining effect, which leads to a single huge heterogeneous cluster containing almost every streamline (Section 5.1). Hence, cluster cohesion as well as separation are small. Chaining has been observed for all data sets and most numbers of clusters. K-means performs better but sill exhibits a rather low value (0.18). The reason is that simply resampling all streamlines to a uniform number of vertices amplifies differences in streamline length and position offset for otherwise very similar lines. This counteracts our streamline similarity measure, which has been tailored to tolerate these differences (Eq. (2)). As a consequence, similar lines are assigned to different clusters. Complete link also achieves a low silhouette width of 0.28. This is likely due to its tendency to break large clusters leading to a low inter-cluster separation between the resulting parts. This effect could be observed on a sample basis. Average link, Ward's method and SC perform equally well and exhibit the highest silhouette widths: $0.42, 0.43, 0.38$.

The silhouette width is biased towards globular clusters [39]. In case of elongated or concave clusters, algorithms correctly identifying them, e.g., single link and SC, may be assigned a lower silhouette width than failing algorithms. Since the cluster shape in streamline space is not clear, the silhouette width must be employed carefully. For fiber tracts, the non-globular nature of clusters has already been acknowledged [44].

Connectivity: Single link clustering by far achieves the best connectivity value due to its proximity measure which strives for a merge with the nearest neighbor. This bias has already been acknowledged in [39]. The second and third best connectivity values are achieved by average link and Ward's method. Complete link exhibits the worst value of all AHC variants. It more often adds similar neighbors of a streamline to another cluster, which may again be due to the breaking of large clusters. This leads to streamlines at the joint cluster border, which have similar neighbors in both clusters. The connectivity of SC is worse than for all AHC variants. However, this is to a great extent caused by the functioning of the algorithms and the way of computing connectivity. The computation adds the highest penalty value if the most similar neighbor is not in the same cluster. This rarely occurs in AHC since each variant starts by locally aggregating the nearest singleton clusters. SC aims at a global optimization and occasionally adds the most similar line to another cluster. A preliminary investigation revealed this phenomenon at the joint border of closely spaced clusters. Due to the bias of connectivity towards the AHC approaches, its usefulness in assessing SC is questionable. The connectivity of k-means is worst for the same reason as for the silhouette width.

Hubert's Γ Statistic: Hubert's Γ Statistic shows a poor result for single link due to the chaining effect (0.04). In the one large cluster, very dissimilar streamlines are grouped together leading to negative correlation values. The performance of k-means is considerably better (0.39) but still worse than for the remaining algorithms since the above-mentioned assignment of similar streamlines to different clusters leads to negative correlation values. Complete link, SC, and Ward's method reach similar results on average $(0.44, 0.45, 0.48)$. The highest value is measured for average link by a rather narrow margin (0.52).

Stability: Single link's stability (0.97) is not expressive since the entire set of streamlines is always grouped in a single cluster. Complete linked achieves the lowest stability (0.59) due to the maximum computation in the proximity measure (Section 5.1). Since random subsamples are drawn from the original data to measure stability, different streamlines are missing each time. While the maximum computation is considerably affected by missing lines, the average and the variance computation in Average link and Ward's method, respectively are less sensitive $(0.79, 0.74)$. SC and k-means achieve the highest meaningful stability values $(0.82, 0.85)$. Both apply a global objective function and are hence, less sensitive to local changes than AHC. However, the stability of k-means is dependent on the number of runs (=10, Section 5.2) and decreases to 0.72 for a single run. Even with a high number of runs, k-means may generate different results if started several times due to the random initialization of cluster centroids. The result of all AHC variants is dependent on the order of the input streamlines. If the proximity measure happens to be equal for two pairs of clusters, the first encountered pair is merged. However, we did not observe this problem.

Summary: Single link is not suitable for clustering blood flow due to the chaining effect which requires dedicated post-processing [10]. Complete link generates better clusters but tends to break large clusters. This has a negative impact

on inter-cluster separation, which is reflected by lower silhouette widths. Further, the clustering results of complete link show a rather low stability. Stability becomes an important issue if the seeding density is varied, e.g., along the ostium, or in interactively sampling a region-of-interest by overlapping seeding regions, e.g., the aneurysmal near-wall region. In both cases, pairs of similar streamlines that survive the modifications should consistently be assigned to a joint cluster. K-means performed particularly poor with respect to the silhouette width and connectivity. Also, the stability of its clusters is less predictable due to the random initialization. Average link, Ward's method, and SC performed equally well except for the connectivity which is however biased towards AHC. An extended evaluation may investigate the overlap of their clustering results to gain further insight into their principles of operation and the data.

Average link's sensitivity to outliers was significantly reduced by our adapted streamline similarity measure (Section 4.2). While the original measure (Eq. (1)) lead to small-sized, outlier-corrupted clusters (<6 streamlines) in each data set, this effect was only observed in three data sets with the new measure. Ward's method and SC proved to be rather insensitive to outliers. Overall, we recommend Average link, Ward's method, and SC for clustering blood flow. Visual blood flow summaries based on each of them are qualitatively evaluated by domain experts in Section 6.4.

6 Visual Summary of Blood Flow

This section is dedicated to the computation of cluster representatives, their aggregation in a visual flow summary, the interaction with the summary, the expert evaluation of the summary, and our development environment. The accompanying supplemental video, which can be found at http://dx.doi.org/10.1109/TVCG.2013.2297914, shows a use case of the flow summary.

6.1 Cluster Representatives

Displaying thousands of streamlines leads to a cluttered visualization hampering particularly the interpretation of inner flow structures (Fig. 4a). Cluster representatives summarize the flow and show these structures (Fig. 4b). In the context of clustering fiber tracts, different types of representatives have been discussed [33]. O'Donnel et al. employ spectral clustering and determine an *embedding-based* representative for each fiber bundle in spectral embedding space (Fig. 2b). The centroid of the bundle's point cloud is computed and the fiber closest to it is chosen. This is feasible due to the high density and number of embedded fibers (up to $25{,}000$ per brain). In our case, the streamline count is often $<3{,}000$. Furthermore, given a non-globular cluster, e.g., banana-shaped, the streamline closest to the cluster centroid may provide a weak representative.

As an alternative computed in the original 3D space, we chose the streamline with the smallest average distance to all other lines of the cluster. While often well representing the clusters, this *distance-based representative* is prone to outlier streamlines due to the outer minimum in the distance measure (Eq. (2)). A short outlier, running very similar to all streamlines in its cluster, is assigned a small distance to all of them. Longer streamlines are more likely to deviate from the other lines in their cluster. Hence, the outlier is a more likely candidate for representative selection.

O'Donnell et al. propose another approach for computing representatives in 3D space [33]. For each cluster of fibers, a local Cartesian grid is aligned with the cluster's axis-aligned bounding box. For each voxel of the grid, the number of fibers that pass through is recorded leading to a density volume. For each fiber, the density is integrated along the line and the result is weighted with the fiber's length. The fiber with the highest value is the *density-based representative*. Several problems occur in transferring this approach to streamlines. The lines in a cluster may follow the same course over a long range but extend beyond either end of this range (Fig. 4c, bottom). No line may exist that faithfully represents the entire cluster. The lines may also differ significantly in length. Furthermore, a few very long lines may exist in helical flow. Hence, we consider only density and for now neglect the weighting with length. Note that length is still inherently considered, since longer lines may accumulate more densities. The primarily density-based representatives well indicate the densest parts of the clusters which often occur in regions of helical or turbulent flow being of high interest. In an initial flow summary and in the remainder of this paper, we employ density-based representatives. However, the user may change the flow

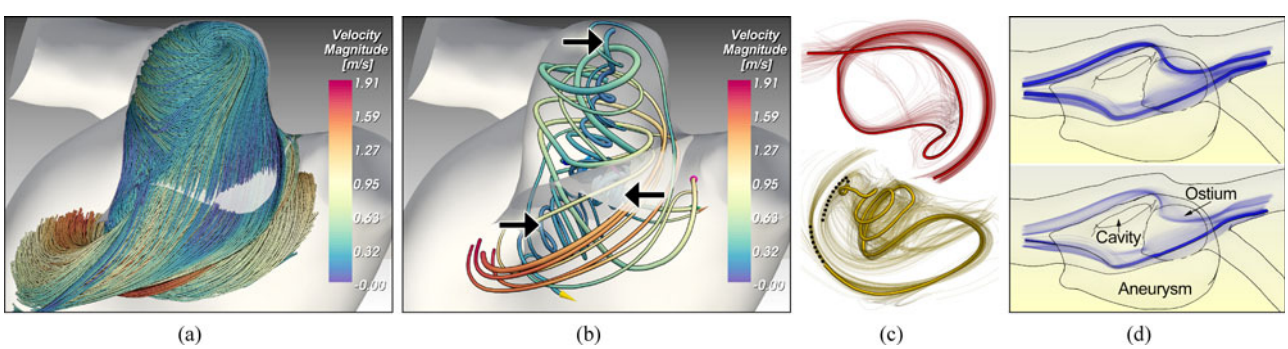

Fig. 4. (a) Full set of streamlines in a side-wall aneurysm. (b) Streamlines in (a) clustered according to geometry. One representative is displayed for each cluster ($n = 9$). A prominent swirl in the center of the aneurysm and laminar helical and complex flow below the ostium (transparent surface) are revealed (top, left, and right arrow). (c) Examples for good (red) and amendable (yellow) representatives. Dots indicate parts of the cluster which are not represented. (d) Flow around a cavity clustered according to local residence time. A selected cluster is visualized by semi-transparent streamlines. Its attribute-based representative indicates only the lower branch (bottom). A representation of cluster shape is obtained by further clustering based on streamline geometry (top).

summary by modifying weights $[0, 1]$, which we assigned to density, length, and distance. For instance, setting the weight of density to zero and the weight of length to one leads to length-based representatives, which may better illustrate the entire extent of the cluster.

If streamlines were clustered according to a streamline attribute, we employ *attribute-based representatives*. For each cluster, the mean of the attribute or of the statistic that has been employed is computed and the line with an attribute value closest to the mean is chosen (Section 4.3). The representative then indicates the clusters attribute range instead of its shape. Since the course of streamlines inside a cluster may be rather heterogeneous, we conduct a further partitioning according to streamline geometry (Fig. 4d).

6.2 Number of Clusters

A crucial question in generating the blood flow summary is how many representatives should be displayed, i.e., how many clusters must be computed? For blood flow data, the "correct" number of clusters is not known. Agglomerative hierarchical clustering is well suited here since the cluster hierarchy may be cut at consecutive levels in order to interactively browse through a range of cluster numbers (Section 5.1). Spectral clustering and k-means require rerunning the algorithm each time. Merging and splitting clusters in AHC occurs locally in space and is hence easier to track visually. However, our practical experience with highly intertwined 3D streamline clusters shows that it is still difficult to grasp the change between consecutive cluster numbers without visual guidance.

We aim at minimizing the workload of physicians by making a "good guess" with respect to the number of clusters. A default number increases the reproducibility of our approach, which is a key requirement for entering clinical routine. Further, it facilitates a more standardized comparison of the flow before and after stenting and it supports a categorization of blood flow patterns. A good guess leads to clusters representing all significantly distinct flow structures—overrepresented structures are tolerable while missing structures are not—and each cluster is homogeneous such that the representative indeed represents all contained streamlines. Translated into clustering language, the inter-cluster separation and the intra-cluster cohesion should be high. We couple the quantitatively best performing streamline clustering techniques, AHC with average link, AHC with Ward's method, and SC (Section 5.4), with state-of-the-art techniques computing the number of clusters k that best satisfies both requirements.

Salvador and Chan propose the *L-method* for computing k in hierarchical clustering algorithms [45]. The method is based on detecting the *knee* in a graph that opposes numbers of clusters and a cluster evaluation metric. Since the location of the knee depends on the shape of the graph which again depends on the number of tested cluster numbers, a full evaluation graph, ranging from two clusters to the number of data elements, is recommended. We compute the full graph based on the evaluation metric suggested in [45]. Zelnik-Manor and Perona propose an algorithm for computing k in SC [42]. The algorithm iterates over a user-defined range $[a, b]$ for k and determines the optimal value. The optimization is based on finding the optimal rotation between the set of the first $k_i, i \in [a, b]$ largest eigenvectors of the Graph Laplacian (Eq. (4)) and the canonical coordinate system. We empirically determined the range $[4, 20]$ for detecting all relevant flow structures in 10 data sets.

6.3 Visualization and Interaction

In the initial blood flow summary, cluster representatives corresponding to the optimal partitioning are shown (Fig. 4b). The user may inspect the suboptimal partitionings by browsing AHC's hierarchy or SC's range $[a, b]$. A representative can be picked causing the corresponding cluster to be displayed. For browsing all clusters, the user may scroll the mouse wheel. If the clustering was based on a streamline attribute, the set of geometry-based representatives per cluster is displayed after picking and during browsing (Section 6.1).

The streamline visualization is embedded in a surface rendering of the vessel wall. The wall is reconstructed from the unstructured grid of the CFD simulation. It is rendered opaque with culled front faces. The opaque back faces prevent a look through the aneurysm on lines in the near-vessel domain. The ostium and the stent surface are integrated. The ostium is rendered highly transparent.

Streamlines are rendered with GPU support as sets of quads and halos are added to improve spatial perception [46] (Fig. 4a). The halo color is either set to black or encodes the cluster ID. The latter is useful to distinguish clusters when the line color is modified according to a streamline attribute. However, our collaborators criticized the interference of halo and line color hampering the readability of the attribute. We initially color all halos in black and optionally allow an encoding of the cluster ID.

For visualizing the representatives, we evaluated stream ribbons and tubes. While ribbons additionally show rotation about the flow axis, color-mapped values are easier to read from tubes during a change of the viewing perspective. Our collaborators rated the readability as more important and hence, we employ tubes. In order to illustrate the flow direction, arrowhead glyphs are attached to the end of each tube pointing in outflow direction. The tube radius encodes the cluster size, i.e., the number of grouped streamlines. Halos are added to the representatives and initially colored in black. While this solves the color interference problem, it hampers visually tracking a tube through the set of highly intertwined representatives. Hence, we offer an optional color encoding of the cluster ID. Alternatively, only the halo of the representative under the pointer is colored according to cluster ID during mouse hover and the other representatives are rendered semi-transparently.

An important aspect is the coloring of streamlines and representatives. In geometry-based clustering, streamline color is modified according to a user-defined attribute. In attribute-based clustering, the statistic that has been employed for clustering is displayed per line, e.g., the maximum or mean of the attribute (Section 4.3). Two approaches are implemented for coloring the representatives: (1) simply copying the attribute values of the corresponding streamline, and (2) averaging the attribute values over all lines in the cluster. If the clustering has been based on streamline geometry, we apply (1) for attributes being defined as a

series of values along each streamline and (2) for single scalar attributes. Note that (1) provides a reasonable approximation of the entire cluster for most flow attributes since their change in value is similar across all streamlines in the cluster due to the common underlying flow pattern. If the clustering has been based on an attribute, we directly apply (2) for single scalar attributes and for a series of values, we average over the statistic that has been employed in computing streamline similarity (Section 4.3).

6.4 Qualitative Evaluation

We let domain experts evaluate blood flow summaries generated by means of the quantitatively best performing streamline clustering techniques (Section 5.4): agglomerative hierarchical clustering with average link, AHC with Ward's method, and spectral clustering. The number of clusters in the summary and hence, the number of representatives, has been computed automatically (Section 6.4). The evaluation is based on three clinical cases together comprising five data sets and representing the prevailing types of aneurysms. One case has been simulated without virtual stenting (Fig. 4a). Two cases have been simulated with and without stenting, one of them with two types of stents in two different positions (Sections 7.1 and 7.2). For the latter case, we considered only the most beneficial type of stent and position. The blood flow summaries were evaluated by two board certified (BC) senior interventional neuroradiologists, a BC senior radiologist with a strong background in aortic aneurysms, two CFD engineers with a strong background in cerebral blood flow (one being coauthor of the paper), and one computer scientist working on experimental 7-Tesla magnetic resonance imaging (MRI) of cerebral blood flow. The CFD engineers and one of the neuroradiologists participated in the virtual intracranial stenting challenges in 2009 and 2010 (Sections 7.1 and 7.2). The case without virtual stenting was stented by the neuroradiologist in real life.

Flow Summary: At first, the experts were asked to familiarize with the original data, i.e., the streamlines. All of them had seen streamline visualizations of blood flow before. However, the two neuroradiologists had no and only limited experience, respectively in interacting with such visualizations, e.g., filtering lines and probing by interactive seeding. The streamlines were visualized as in Fig. 4a. The experts could filter lines by thresholding their average distance to the vessel wall. This offered browsing through the lines from the vessel wall to the center in order to grasp the path of the flow through the near-vessel domain (Fig. 1b) and to detect characteristic flow structures, such as swirls. The experts were asked to sketch the flow path and annotate all structures that they consider to be relevant in a drawing of the aneurysmal silhouette.

Then, the flow summaries based on the three clustering algorithms were presented in a random, blinded side-by-side arrangement. In addition, a *control summary* was generated and mixed in to eliminate coincidence. This summary was generated based on a random number k of clusters, with k being in the range of the numbers computed for the three algorithms. Cluster size, the assignment of streamlines to clusters, and the selection of cluster representatives were also randomized. The experts were asked to rate each flow summary. Zero points were given if the sketched flow was in no way represented by the summary, one point was given if it was partially represented and two points in case of full representation. Finally, the experts should check whether the summary reveals other important patterns than they had discovered. Additional comments were recorded during the evaluation. The overall time exposure for the experts was $\approx$60 minutes.

TABLE 2
Average Expert Ratings of Blood Flow Summaries
($\in \{0, 1, 2\}$, $2 = \text{Best}$)

Algorithm	Datasets (NVS=Not Virtually Stented, V09/V10=Virtual Intracranial Stenting Challenge 2009/10, S=SILK, R=right posterior cerebral artery)					
	NVS	*V09*	*V09S*	*V10*	$V10S_R$	$\varnothing$
SC	2.0	2.0	1.5	2.0	2.0	1.9
AHC_avg	1.5	1.8	2.0	2.0	1.2	1.7
AHC_ward	1.8	1.8	0.8	1.8	1.2	1.5
RAND	1.0	1.0	0.0	0.7	1.0	0.7

Comparison of spectral clustering, agglomerative hierarchical clustering with average link (avg) and Ward's method, and random generation (RAND).

The results of the evaluation are summarized in Table 2. SC consistently achieves the best results. Except for one data set, its flow summaries fully represent the flow sketched by the experts. For this specific data set, half of the participants considered a swirl as "not really visible" (one point) while the other half considered it to be "slightly indicated" (two points). AHC with average link and with Ward's method show the second and third best results, respectively. However, Ward's method never achieves the full score on average for none of the data sets. The control summary (RAND) performs significantly worse than the rest, which confirms that the other summaries indeed provide non-random, meaningful insight. In 33 flow summaries out of 90 (five data sets times six participants times three algorithms, excluding RAND), the experts detected more interesting flow patterns than they had discovered during streamline filtering further indicating the summary's benefit. The 33 summaries were generated in equal shares by the algorithms thus not indicating a unique feature.

Number of Clusters: The CFD engineers and the computer scientist were given an extra task before the assessment of the flow summaries. This time-consuming task did not fit into the tight schedule of the physicians since it extended the evaluation time to 90–120 minutes. In a sequence, the flow summaries based on the range of possible numbers of clusters $[4, 20]$ were presented and the experts were asked to select the number k_{sel} that fully represents their sketched flow, possibly shows more important flow structures, and is still clearly readable. To reduce time exposure, each expert assessed each data set only based on one alternately chosen algorithm A with the control summary being left out (3 experts times five algorithms results in 15 ratings). After k_{sel} had been determined, the experts were asked to rate the flow summaries as explained above. Afterwards, the summary corresponding to A was pointed out and the expert was asked to compare the associated computed number of clusters k_{cmp} to k_{sel}.

For SC, k_{sel} was preferred once over k_{cmp}, namely for the only data set for which SC's flow summary did not achieve the full score on average (Table 2, $V09S$). For both AHC with average link and AHC with Ward's method, k_{sel} was preferred three times over k_{cmp} since important flow structures were missing based on k_{cmp}. The remaining eight comparisons assessed k_{cmp} as appropriate for generating an uncluttered summary, which is complete with respect to characteristic flow structures. In five (of eight) comparisons, these structures were overrepresented ($k_{cmp} > k_{sel}$) but still clearly visible. In the remaining three comparisons, k_{sel} was higher than k_{cmp} because one specific swirl was seen based on both but even more clearly based on k_{sel}.

In conclusion, the blood flow summaries based on SC have achieved the best evaluation results by a narrow margin. The applied clustering algorithm, the number of clusters, and the type of representative effect the success of the summary. Hence, we recommend and employ in the remainder SC, its associated technique for computing a reliable number of clusters, and density-based representatives (Sections 5.3, 6.1, 6.2). Since k_{cmp} was assessed as inappropriate in one case of SC, we offer interactively browsing the range of possible cluster numbers $[4, 20]$ starting from k_{cmp}.

Anecdotal Feedback: All experts agreed that the flow summary is much faster to interpret than the entire set of streamlines and reveals flow features which are hidden inside the streamline clutter. They appreciated the workload reduction by avoiding the tedious iterative procedure of selectively seeding and/or filtering streamlines. Displaying streamline clusters on demand was rated as very valuable to get an impression of the spatial region that is represented by a cluster representative. Supporting the visual tracking of individual representatives by coloring the halo of the representative under the mouse pointer was preferred over temporarily modifying the halo color of all representatives according to cluster ID (Section 6.1). The physicians agreed that the comparison of flow before and after stenting is greatly simplified by the flow summaries in Figs. 5 and 7.

6.5 Design Lessons

We carefully designed the flow summary in a tight feedback loop with our collaborators. The design lessons learned help other visualization practitioners working with similar data.

(1) Restrict the clustering domain to the region-of-interest. We restrict it to the aneurysm and the near-vessel domain. Otherwise, long sections of straight in- and outflow would lead to high streamline similarities while differences inside the aneurysm would have less impact (Section 4.1). (2) Choose a similarity measure that is less sensitive to streamline length if the course of streamlines is the primary concern. (3) Provide a good initial guess of the number of clusters since visually tracking the changes while browsing through different numbers of clusters is a tedious task especially for highly intertwined streamlines. (4) Use tubes as cluster representatives instead of ribbons if the readability of attribute values is crucial. (5) Add halos to streamlines and representatives in order to enhance their spatial perception. (6) Use black as halo color to avoid visual interference with color-coded streamline attributes. (7) Support visual tracking of tubes through a set of intertwined representatives by assigning a striking color to the halo of the representative under the mouse pointer. (8) Allow the user to see the original clusters since the representatives well encode the general course of the contained streamlines but fail in illustrating the cluster extent. (9) Encode the direction of the flow, e.g., by arrowhead glyphs. (10) Attribute-based clustering may require the computation of several representatives per cluster since the streamlines in a cluster may be quite heterogeneous with respect to their geometry (Figs. 8b and 8d).

6.6 Development Environment

The clustering algorithms, the similarity measures and the computation of cluster representatives are implemented in Matlab (MathWorks, Natick, MA). Source code for local scaling and determining the number of clusters is provided by Zelnik-Manor and Perona [47]. All Matlab code is exported as a shared library and accessed from custom C++ code. The three categories of streamline attributes are computed using (1) ANSYS Fluent 12 and ParaView, (2) the vascular modeling toolkit (www.vmtk.org), and (3) custom C++ code (Section 4.3). The visualization is implemented in C++ and the Visualization Toolkit (Kitware, Inc., Clifton Park, NY).

7 Application

We applied our approach to data of the virtual intracranial stenting challenges in 2009 and 2010 [48]. Please consider the following advices when reading the figures of this section. The color scales refer to the representatives, not their halos. The annotated range of values is based on the entire set of streamlines. Halo colors must not be employed for establishing correspondence between clusters in different figures or figure parts. They are assigned independently to each clustering result and simplify the visual tracking of representatives in a non-interactive display.

7.1 Virtual Intracranial Stenting Challenge 2009

For the VISC 2009, teams were invited to compete in predicting stenting success based on simulated hemodynamic data. Two cases and a model description of the flow diverting *SILK* stent (Balt, Montmorency, France) were provided. Due to space restrictions, we only discuss the first case with a saccular side-wall aneurysm located at a bifurcation (Fig. 5a). A rare anatomical variant is the cavity (*fenestration*) behind the aneurysm. Our medical collaborators suggested placing the stent in the right artery and circumventing the aneurysm to the left. The stent geometry was modeled in a CAD program and manually fitted to the vessel wall. The hemodynamic data generation resulted in volume meshes with 4.3 and 4.6 (with stent) million tetrahedral elements (Section 2.3). The meshes constituted the input for streamline generation (Section 4.1).

The resulting lines have been clustered based on geometry (Section 4.2). The flow summaries are displayed in Figs. 5b and 5c. A higher number of clusters can be observed in the untreated aneurysm indicating a more complex flow pattern (Fig. 5b). After stenting, the flow is less complex which decreases the risk of aneurysm rupture [5].

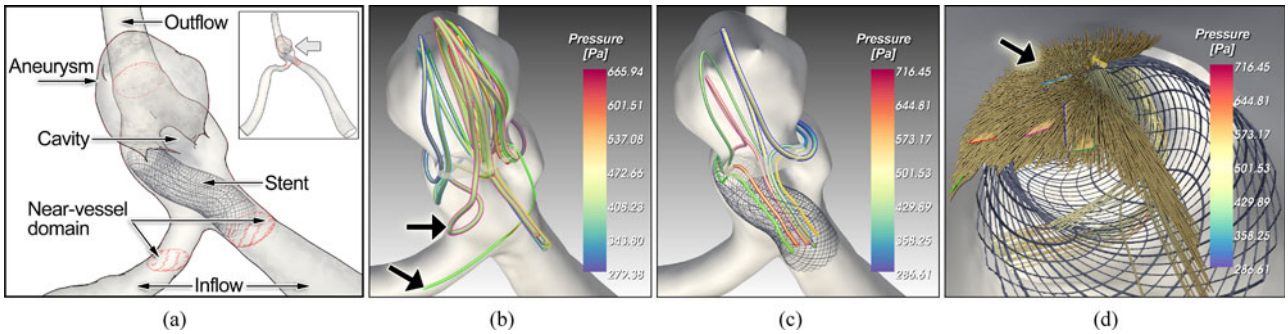

Fig. 5. VISC 2009. (a) Virtual stent placement, morphological features, subdivision of vascular domain (red circles) and flow conditions. (b,c) Streamlines clustered based on geometry (b) before and (c) after stenting. Arrows point at interesting differences, e.g., reflux (upper arrow). Flow from the right artery is not completely diverted (c). (d) Inside the artery. Flow bypassing the stent (arrow) reveals a gap between stent and vessel wall.

In the stented configuration, flow arriving from the right artery is not completely diverted but still enters the aneurysm (Fig. 5c). A closer look from inside the vessel at the location where this flow enters the stent reveals that the stent model does not perfectly adhere to the vessel wall (Fig. 5d). A considerable gap exists through which flow with high pressure is bypassing the stent. A neuroradiologist commented that such gaps indeed occur in real stenting due to a sharp bending of the vessel. Their prediction would be of great value. The flow that travels through the virtual stent, exits the stent at its aneurysm-near inflection point and enters the aneurysm (Fig. 5c). This may be mitigated by a higher general or local mesh density. Before stenting, reflux is observed below the ostium (Fig. 5b, top arrow). Furthermore, flow is entering the aneurysm from the left artery with high pressure (Fig. 5b, bottom arrow). After stenting, this flow is obstructed by the stent and circumvents the aneurysm. This is a convenient side effect of diverting the flow arriving from the right branch.

A comparison of the aneurysmal wall-shear stress before and after stenting revealed lower values in the latter which indicates a benefit. We investigate the near-wall flow by clustering the streamlines based on their mean distance to the aneurysm wall (Section 4.3). The results before and after stenting are compared in Fig. 6. To support a visual comparison, the color mapping and the radius scaling of the representatives after stenting are applied uniformly to both configurations. The comparison shows that more flow hits the wall and is traveling through the near-wall region before stenting. This is in accordance with the higher WSS [49]. After stenting, a considerable amount of the flow barely enters the aneurysm (thick blue tube in Fig. 6b). Note that attribute-based representatives have been applied well indicating a cluster's range of attribute values (Section 6.1).

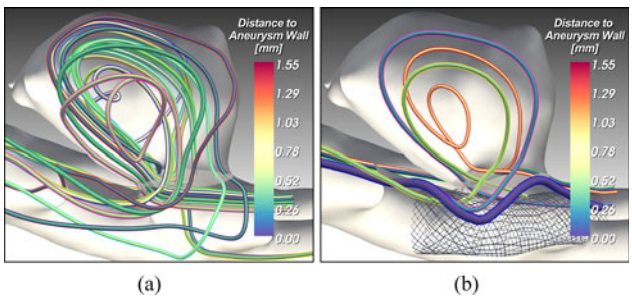

Fig. 6. Clustering streamlines according to their mean distance to the aneurysm wall before (a) and after virtual stenting (b). Stenting reduces near-wall flow.

7.2 Virtual Intracranial Stenting Challenge 2010

For the VISC in 2010, research teams were invited to find the optimal placement of a stent in treating a basilar tip aneurysm (Fig. 7a). We considered two types of stents and two different positions, both covering the end of the basilar artery and then extending to the beginning of the left and the right posterior cerebral artery (LPCA/RPCA), respectively. We restrict our discussion to the most beneficial type of stent (SILK). The hemodynamic data of the two stented configurations and the untreated case has been generated as described in Section 2.3. The biggest tetrahedral mesh consists of 13.5 million elements (including stent). The stent geometry was modeled in a CAD program. Learning from the issues of a manual stent deployment (Section 7.1), we applied an automatic wall-tight deployment using polyharmonic splines for free-form deformation [15].

For the detection of flow structures in the untreated aneurysm and in the two stented configurations, the near-vessel domain is specified (Fig. 7a) and the data is cropped. Then, streamlines are seeded at the ostium and clustered based on geometry. Cluster representatives are displayed and colored according to local residence time (RT, Section 4.3). The color scale has been set for all configurations to mapping the range of RT in the untreated configuration (Figs. 7a, 7b, 7c and 7d). Thus, regions of prolonged RT after stenting can be easily spotted. Before we focus on RT, we study the detected flow structures.

In (Figs. 7b, 7c and 7d), representatives indicating a major difference between the flow patterns are rendered opaque. Before and after stenting along the LPCA, parts of the flow enter the aneurysm and after a swirling motion inside, exit via the RPCA (Figs. 7b and 7c). Stenting along the RPCA considerably alters the flow pattern and generates a double helical swirl in the center of the aneurysm. A closer look at the highlighted representative (s) of each configuration revealed that they always represent those clusters with the highest RT values on average. Comparing their coloring indicates that SILK stenting along the RPCA causes the most prolonged RT and hence represents the preferred strategy (Fig. 7d). Further evidence is given by plotting the percentage of streamlines

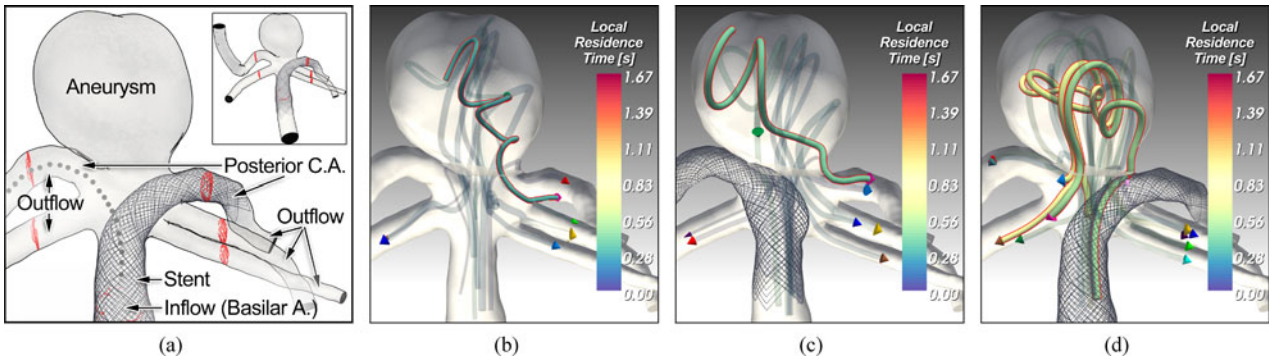

Fig. 7. VISC 2010. (a) Virtual stent placement, morphological features, subdivision of the vascular domain (red circles) and flow conditions. (b-d) Clustering of streamlines based on geometry before (b) and after stenting (c,d). Representatives indicating a major difference between the flow patterns are rendered opaque. While a "simple" swirl is characteristic for the first two patterns (b,c), a double helical swirl is observed in the third one (d).

over discrete RT values (Fig. 8a) and by Janiga et al. [15] who report the most prolonged *turnover time* for this configuration. The turnover time is a global scalar measure which is proportional to RT and both characterize intra-aneurysmal flow stasis [50]. In the following, we focus on stenting along the RPCA.

In order to investigate RT more locally, the streamlines of the stented configuration have been clustered based on it. The cluster with the highest RT values is shown in Fig. 8b. Its streamlines are rendered semi-transparent such that the inner swirl is easier to perceive. Flow enters the aneurysm, is attracted by opposing wall parts, converges in a swirl in the center, and leaves the aneurysm (the swirling motion is also indicated in Fig. 7d). Since correspondences between a low WSS and thrombosis development as well as between a high RT and thrombosis development are known [4], the cluster has been further investigated in the context of WSS (Fig. 8c).

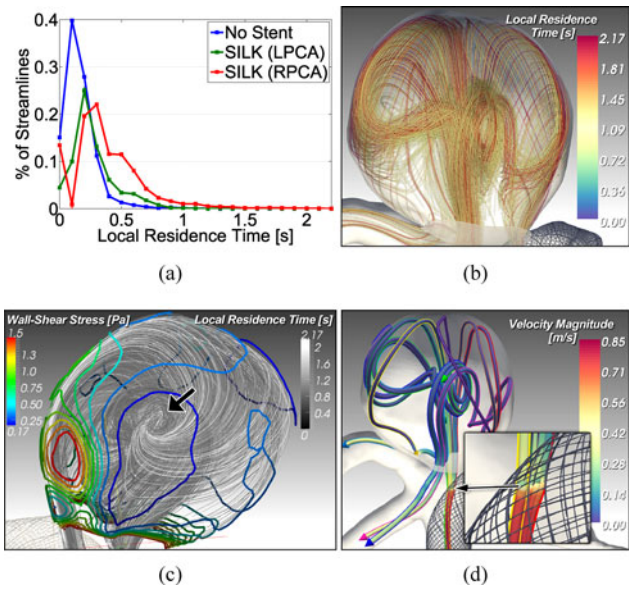

Fig. 8. VISC 2010. (a) Comparison of local residence times before and after stenting along the left and right posterior cerebral artery (LPCA, RPCA). The percent of streamlines is plotted over discrete RT values. Stenting causes prolonged RT. (b) Cluster with the highest average RT in RPCA stenting. (c) Investigating this cluster in the context of wall-shear stress (iso-contours). (d) Partitioning the cluster based on streamline geometry. Flow is strongly decelerated at the stent wires (inset).

WSS is mapped to the surface of the aneurysm and visualized by contour lines. In agreement with [4], a value of 1.5 is chosen as the upper limit for color mapping. Values above are clamped to this limit. As can be observed, low WSS values occur in a large region where the flow streaks the wall (arrow). The same is true for the opposite side of the wall. It should be further investigated whether these regions are potential candidates for thrombosis initiation and whether a double helical swirl particularly encourages flow stasis.

It was shown in [15] that the SILK stent diverts a considerable amount of blood. However, parts of the blood flow exit the wired mesh and enter the aneurysm. In Fig. 8d, the cluster with the highest values of RT is refined by a clustering based on geometry and the new representatives are colored according to velocity magnitude. The structure of the swirl is easier to perceive as compared to Fig. 8b. Furthermore, it can be observed that the flow exiting the stent is strongly decelerated at its wires thus leading to a prolonged RT and a slow inflow (strong red to green jump in the inset of Fig. 8d).

7.3 Performance

This section reports on the performance of our approach. The focus is on computation time since memory consumption is not critical. The time is dependent on the number of streamlines and their number of vertices (columns 2-3 in Table 3). While the first varies with the sampling density of the ostium, the latter depends on the streamline length and integration step size (Section 4.1). We measured the computation time of spectral clustering and of the visualization. In clustering, we differentiated between the computation of

TABLE 3
Data Set Characteristics and Timings $[s]$ of Geometry-Based Clustering and Visualization

Dataset	#Streamlines	#Vertices (⌀)	Distance Matrix	Clustering	Visualization
$V09$	2254	505	4547	34.7	9.5
$V09S$	2207	249	732	29.9	4.4
$V10$	2929	265	1567	57.2	6.2
$V10N_R$	2923	275	1581	55.9	6.8
$V10N_L$	1153	283	256	12.8	4.1
$V10S_R$	2891	234	1128	51.8	7.4
$V10S_L$	1138	212	142	14.1	3.7

V09/V10 = Virtual Intracranial Stenting Challenge 2009/2010, S = SILK stent, N = Neuroform stent, L/R = left/right posterior cerebral artery.

the distance matrix and the actual clustering. The latter also comprises the determination of cluster representatives. In distance matrix computation, we focused on geometry-based distances since attribute-based distances are much faster to compute. The timings were taken on a 3.07 GHz Intel 8-core PC with 8 GB RAM and a 64 bit Windows operating system (Table 3).

As expected, the computation of the distance matrix represents the bottleneck. However, the matrix can be reused for different clustering settings. In attribute-based clustering, the time for computing the matrix depends on the applied statistic (Section 4.3). For simple statistics such as min/max, the computation is two orders of magnitude faster than in geometry-based clustering. The timings for the clustering itself are in the range of seconds. The most time-consuming part of the visualization is the geometry computation for the GPU-based streamline rendering.

8 Summary and Discussion

We presented an approach for reducing visual clutter in streamline visualizations of simulated blood flow. The approach is based on clustering streamlines and computing cluster representatives, which are compiled into a flow summary. To determine the most appropriate clustering algorithm, we carried out a quantitative evaluation of spectral clustering, four variants of agglomerative hierarchical clustering, and k-means. Based on cluster validity measures, we identified SC and AHC with average link and Ward's method, respectively as superior. In an expert evaluation of blood flow summaries generated by these algorithms, SC achieved the best ratings by a narrow margin. Its summaries are complete with respect to the relevant flow structures in most cases. In a tight feedback loop with our collaborators, we carefully designed the flow summary. The design lessons learned help scientists, e.g., in exploring flow in other vascular pathologies.

The computation of the summary is fully automatic. Only a range for the number of clusters that possibly exist in the data set must be provided. The optimal number is identified automatically. We empirically determined a range of $[4, 20]$ for detecting all relevant flow structures in ten data sets. The time needed for the clustering and the visualization is within the range of minutes. Compared to the Computational Fluid Dynamics simulation, which takes hours, it is of little consequence with respect to a possible therapeutic workflow.

Results from CFD simulations are not yet part of the clinical decision pipeline although they can be generated within a clinically acceptable time frame for planning an intervention. Neuroradiologists have little experience in investigating flow data. Our flow summary simplifies the access to the data, it is easier to read than full streamline visualizations, and it contributes to the communication between CFD engineers and physicians. The latter is of crucial importance in understanding "How stent properties affect flow patterns?", "How the change in flow patterns after stenting is related to treatment success?", and "How flow patterns are related to the risk of aneurysm rupture and the development of thrombosis?". Once these questions can be answered, stenting may not be planned solely based on the coverage of the aneurysm neck by the stent, but also based on CFD results and the flow summary. The concept of the summary can be readily transferred to (virtual) coiling. However, the joint visualization of coils and cluster representatives will cause serious occlusion problems.

The success of virtual stenting is so far evaluated based on global measures, e.g., the turnover time. However, if a certain stented configuration does not indicate a benefit for the patient, global measures fail to explain why. We cluster streamlines also based on locally derived domain-specific attributes, e.g., the distance to the aneurysm wall and the local residence time. The latter was considered a useful extension to the turnover time. Clusters with a high residence time may forecast locations of thrombosis initiation.

A limitation of our approach is that a few cluster representatives do not capture the entire structure of their cluster. They faithfully represent its densest part but fail to represent all parts in the in- and outflow regions of the near-vessel domain. Hence, the clusters itself should also be inspected.

Acknowledgments

The authors would like to thank Cordula Scherlach, Oliver Beuing, Martin Skalej (University Hospital Magdeburg, Germany), Uta Preim (Community Hospital Magdeburg), Daniel Stucht (Institute of Biomedical Magnetic Resonance, University Magdeburg), and Philipp Berg (Institute of Fluid Mechanics and Thermodynamics, University Magdeburg) for participating in the evaluation and providing valuable feedback. They are grateful to Christian Rössl and Max Zimmermann (University Magdeburg) for discussions on clustering issues.

References

[1] R.D. Brown, J. Huston, R. Hornung, T. Foroud, D.F. Kallmes, D. Kleindorfer, I. Meissner, D. Woo, L. Sauerbeck, and J. Broderick, "Screening for Brain Aneurysm in the Familial Intracranial Aneurysm Study: Frequency and Predictors of Lesion Detection," *Neurosurgery: Pediatrics*, vol. 108, no. 6, pp. 1132-1138, 2008.

[2] I. Larrabide, M.L. Aguilar, H.G. Morales, A.J. Geers, Z. Kulcsár, D. Rüfenacht, and A.F. Frangi, "Intra-Aneurysmal Pressure and Flow Changes Induced by Flow Diverters: Relation to Aneurysm Size and Shape," *Am. J. Neuroradiology*, vol. 34, no. 4, pp. 816-822, 2012.

[3] J. Cebral, F. Mut, M. Raschi, E. Scrivano, R. Ceratto, P. Lylyk, and C. Putman, "Aneurysm Rupture Following Treatment with Flow-Diverting Stents: Computational Hemodynamics Analysis of Treatment," *Am. J. Neuroradiology*, vol. 32, no. 1, pp. 27-33, 2011.

[4] V. Rayz, L. Boussel, L. Ge, J. Leach, A. Martin, M. Lawton, C. McCulloch, and D. Saloner, "Flow Residence Time and Regions of Intraluminal Thrombus Deposition in Intracranial Aneurysms," *Annals of Biomedical Eng.*, vol. 38, no. 10, pp. 3058-69, 2010.

[5] J.R. Cebral, F. Mut, J. Weir, and C.M. Putman, "Association of Hemodynamic Characteristics and Cerebral Aneurysm Rupture," *Am. J. Neuroradiology*, vol. 32, no. 2, pp. 264-270, 2011.

[6] S. Appanaboyina, F. Mut, R. Löhner, C. Putman, and J. Cebral, "Simulation of Intracranial Aneurysm Stenting: Techniques and Challenges," *Computer Methods in Applied Mechanics and Eng.*, vol. 198, no. 45-46, pp. 3567-3582, 2009.

[7] M. Kim, E.I. Levy, H. Meng, and L.N. Hopkins, "Quantification of Hemodynamic Changes Induced by Virtual Placement of Multiple Stents across a Wide-Necked Basilar Trunk Aneurysm," *Neurosurgery*, vol. 61, no. 6, pp. 1305-1313, 2007.

[8] A.R. Mantha, G. Benndorf, A. Hernandez, and R.W. Metcalfe, "Stability of Pulsatile Blood Flow at the Ostium of Cerebral Aneurysms," *J. Biomechanics*, vol. 42, no. 8, pp. 1081-1087, May 2009.

[9] C.-K. Chen, S. Yan, H. Yu, N. Max, and K.-L. Ma, "An Illustrative Visualization Framework for 3D Vector Fields," *Computer Graphics Forum*, vol. 30, no. 7, pp. 1941-1951, 2011.

[10] H. Yu, C. Wang, C.-K. Shene, and J.H. Chen, "Hierarchical Streamline Bundles," *IEEE Trans. Visualization and Computer Graphics*, vol. 18, no. 8, pp. 1353-1367, Aug. 2012.

[11] C. Rössl and H. Theisel, "Streamline Embedding for 3D Vector Field Exploration," *IEEE Trans. Visualization and Computer Graphics*, vol. 18, no. 3, pp. 407-420, Mar. 2012.

[12] T. McLoughlin, M.W. Jones, R.S. Laramee, R. Malki, I. Masters, and C.D. Hansen, "Similarity Measures for Enhancing Interactive Streamline Seeding," *IEEE Trans. Visualization and Computer Graphics*, vol. 19, no. 8, pp. 1342-53, Aug. 2013.

[13] M. Forsting and I. Wanke, *Intracranial Vascular Malformations and Aneurysms: From Diagnostic Work-Up to Endovascular Therapy*, (series Medical Radiology). Springer, 2008.

[14] N. Andaluz and M. Zuccarello, "Recent Trends in the Treatment of Cerebral Aneurysms: Analysis of a Nationwide Inpatient Database," *Neurosurgery: Pediatrics*, vol. 108, no. 6, pp. 1163-1169, 2008.

[15] G. Janiga, C. Rössl, M. Skalej, and D. Thévenin, "Realistic Virtual Intracranial Stenting and Computational Fluid Dynamics for Treatment Analysis," *J. Biomechanics*, vol. 46, no. 1, pp. 7-12, 2012.

[16] S. Seshadhri, G. Janiga, O. Beuing, M. Skalej, and D. Thévenin, "Impact of Stents and Flow Diverters on Hemodynamics in Idealized Aneurysm Models," *J. Biomechanical Eng.*, vol. 133, no. 7, p. 071005, 2011.

[17] R. Gasteiger, M. Neugebauer, O. Beuing, and B. Preim, "The FLOWLENS: A Focus-and-Context Visualization Approach for Exploration of Blood Flow in Cerebral Aneurysms," *IEEE Trans. Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2183-2192, Dec. 2011.

[18] J. Schöberl, "NETGEN An Advancing Front 2D/3D-Mesh Generator Based on Abstract Rules," *Computing and Visualization in Science*, vol. 1, pp. 41-52, 1997.

[19] M. Neugebauer, G. Janiga, O. Beuing, M. Skalej, and B. Preim, "Anatomy-Guided Multi-Level Exploration of Blood Flow in Cerebral Aneurysms," *Computer Graphics Forum*, vol. 30, no. 3, pp. 1041-1050, 2011.

[20] J.R. Cebral, F. Mut, J. Weir, and C.M. Putman, "Quantitative Characterization of the Hemodynamic Environment in Ruptured and Unruptured Brain Aneurysms," *Am. J. Neuroradiology*, vol. 32, no. 1, pp. 145-151, 2011.

[21] F.H. Post, B. Vrolijk, H. Hauser, R.S. Laramee, and H. Doleisch, "The State of the Art in Flow Visualisation: Feature Extraction and Tracking," *Computer Graphics Forum*, vol. 22, no. 4, pp. 775-792, 2003.

[22] T. Salzbrunn, H. Jänicke, T. Wischgoll, and G. Scheuermann, "The State of the Art in Flow Visualization: Partition-Based Techniques," *Proc. Simulation and Visualization*, pp. 75-92, 2008.

[23] R.F.P. van Pelt, S.S.A.M. Jacobs, B.M. ter Haar Romeny, and A. Vilanova, "Visualization of 4D Blood-Flow Fields by Spatiotemporal Hierarchical Clustering," *Computer Graphics Forum*, vol. 31, no. 3pt2, pp. 1065-1074, 2012.

[24] R.S. Laramee, H. Hauser, L. Zhao, and F.H. Post, "Topology-Based Flow Visualization, the State of the Art," *Topology-Based Methods in Visualization*, H. Hauser, H. Hagen, and H. Theisel, eds., (series Mathematics and Visualization), pp. 1-19, Springer, 2007.

[25] A. Pobitzer, R. Peikert, R. Fuchs, B. Schindler, A. Kuhn, H. Theisel, K. Matković, and H. Hauser, "On the Way Towards Topology-Based Visualization of Unsteady Flow—the State of the Art," *State of the Art Reports*, H. Hauser and E. Reinhard, eds., pp. 137-154, Eurographics Association, 2010.

[26] T. Salzbrunn and G. Scheuermann, "Streamline Predicates," *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1601-1612, Nov./Dec. 2006.

[27] S. Born, M. Pfeifle, M. Markl, and G. Scheuermann, "Visual 4D MRI Blood Flow Analysis with Line Predicates," *Proc. IEEE Pacific Visualization Symp.*, pp. 105-112, 2012.

[28] K. Shi, H. Theisel, H. Hauser, T. Weinkauf, K. Matković, H.-C. Hege, and H.-P. Seidel, "Path Line Attributes—An Information Visualization Approach to Analyzing the Dynamic Behavior of 3D Time-Dependent Flow Fields," *Topology-Based Methods in Visualization II*, (series Mathematics and Visualization), pp. 75-88, Springer, 2009.

[29] A. Pobitzer, A. Lež, K. Matković, and H. Hauser, "A Statistics-Based Dimension Reduction of the Space of Path Line Attributes for Interactive Visual Flow Analysis," *Proc. IEEE Pacific Visualization Symp.*, pp. 113-120, 2012.

[30] R. van Pelt, J. Olivan Bescos, M. Breeuwer, R. Clough, M. Groller, B. ter Haar Romeny, and A. Vilanova, "Interactive Virtual Probing of 4D MRI Blood-Flow," *IEEE Trans. Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2153-2162, Dec. 2011.

[31] R. Gasteiger, D.J. Lehmann, R. van Pelt, G. Janiga, O. Beuing, A. Vilanova, H. Theisel, and B. Preim, "Automatic Detection and Visualization of Qualitative Hemodynamic Characteristics in Cerebral Aneurysms," *IEEE Trans. Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2178-2187, Dec. 2012.

[32] L.J. O'Donnell and C.-F. Westin, "Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas," *IEEE Trans. Medical Imaging*, vol. 26, no. 11, pp. 1562-1575, Nov. 2007.

[33] L. O'Donnell, A.J. Golby, and C.-F. Westin, "Tract-Based Morphometry for White Matter Group Analysis," *NeuroImage*, vol. 45, pp. 832-844, 2009.

[34] B. Moberts, A. Vilanova, and J. van Wijk, "Evaluation of Fiber Clustering Methods for Diffusion Tensor Imaging," *Proc. IEEE Visualization*, pp. 65-72, 2005.

[35] S. Zhang, S. Correia, and D. Laidlaw, "Identifying White-Matter Fiber Bundles in DTI Data Using an Automated Proximity-Based Fiber-Clustering Method," *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 5, pp. 1044-1053, Sept./Oct. 2008.

[36] I. Corouge, S. Gouttard, and G. Gerig, "Towards a Shape Model of White Matter Fiber Bundles Using Diffusion Tensor MRI," *Proc. IEEE Int'l Symp. Biomedical Imaging: Nano to Micro (ISBI)*, pp. 344-347, 2004.

[37] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison Wesley, 2005.

[38] U. von Luxburg, "A Tutorial on Spectral Clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395-416, 2007.

[39] J. Handl, J. Knowles, and D.B. Kell, "Computational Cluster Validation in Post-Genomic Data Analysis," *Bioinformatics*, vol. 21, no. 15, pp. 3201-3212, Aug. 2005.

[40] L. Ferreira and D.B. Hitchcock, "A Comparison of Hierarchical Methods for Clustering Functional Data," *Comm. in Statistics—Simulation and Computation*, vol. 38, no. 9, pp. 1925-1949, 2009.

[41] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, Aug. 2000.

[42] L. Zelnik-Manor and P. Perona, "Self-Tuning Spectral Clustering," *Proc. 18th Ann. Conf. Advances in Neural Information Processing Systems 17*, pp. 1601-1608, 2004.

[43] L. Hubert and P. Arabie, "Comparing Partitions," *J. Classification*, vol. 2, no. 1, pp. 193-218, 1985.

[44] J. Klein, P. Bittihn, P. Ledochowitsch, H.K. Hahn, O. Konrad, J. Rexilius, and H.-O. Peitgen, "Grid-Based Spectral Fiber Clustering," *Proc. SPIE Medical Imaging: Visualization and Image-Guided Procedures*, vol. 6509, pp. 65091E–65091E–10, 2007.

[45] S. Salvador and P. Chan, "Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms," *Proc. Int'l Conf. Tools with Artificial Intelligence (ICTAI)*, pp. 576-584, 2004.

[46] M.H. Everts, H. Bekker, J.B. Roerdink, and T. Isenberg, "Depth-Dependent Halos: Illustrative Rendering of Dense Line Data," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1299-1306, Nov./Dec. 2009.

[47] L. Zelnik-Manor and P. Perona, "Self-Tuning Spectral Clustering Code," www.vision.caltech.edu/lihi/Demos/SelfTuningClustering.html, 2014.

[48] http://www.ics-meeting.net/ics_visc_past_.html, 2014.

[49] G. Janiga, O. Beuing, S. Seshadhri, M. Neugebauer, R. Gasteiger, B. Preim, G. Rose, M. Skalej, and D. Thévenin, "Virtual Stenting Using Real Patient Data," *Proc. Int'l Conf. Fluid Flow Technologies*, pp. 111-117, 2009.

[50] M. Kim, D.B. Taulbee, M. Tremmel, and H. Meng, "Comparison of Two Stents in Modifying Cerebral Aneurysm Hemodynamics," *Annals of Biomedical Eng.*, vol. 36, no. 5, pp. 726-41, 2008.

Steffen Oeltze is a Postdoctoral Research Fellow at the Computer Science Department at the University of Magdeburg, Germany. In 2004, he received the diploma in Computational Visualistics and in 2010, a Ph.D. in Computer Science from the University of Magdeburg. His research interests are in the visual analysis and exploration of medical, biological, and epidemiological data.

Dirk J. Lehmann is a Postdoctoral Research Fellow at the Computer Science Department at the University of Magdeburg, Germany. In 2008, he received the M.Sc. in Computational Visualistics and in 2012, a Ph.D. in Computer Science from the University of Magdeburg. His research interests focus on flow and volume visualization as well as information visualization and visual analytics.

Alexander Kuhn is a Postdoctoral Research Fellow at the Department of Visualization and Data Analysis at the Zuse Insitute Berlin (ZIB), Germany. In 2009, he received the M.Sc. in Computational Visualistics and in 2012, a Ph.D. in Computer Science from the University of Magdeburg. His research interests are in the visualization, segmentation, and analysis of flow data.

Gábor Janiga is researcher (Privatdozent) at the Laboratory of Fluid Dynamics and Technical Flows at the University of Magdeburg, Germany. In 1998, he received his diploma and in 2002 a Ph.D. from the University of Miskolc (Hungary), and a habilitation (venia legendi) in 2011 from the University of Magdeburg. His research interests are in the simulation of blood flow in vascular pathologies and in virtual stenting.

Holger Theisel is professor for Visual Computing at the Computer Science Department at the University of Magdeburg, Germany. In 1994, he received the diploma in Computer Science, in 1996 a Ph.D. in Computer Science, and a habilitation (venia legendi) in 2001 from the University of Rostock. His research interests focus on flow and volume visualization as well as on CAGD, geometry processing and information visualization.

Bernhard Preim is professor for visualization at the Computer Science Department at the University of Magdeburg, Germany. In 1994, he received the diploma in Computer Science and in 1998 a Ph.D. in Computer Science from the University of Magdeburg, and a habilitation (venia legendi) in 2002 from the University of Bremen. His research interests are in medical visualization and applications in diagnosis and treatment.