DOI: 10.1111/j.1467-8659.2012.03100.x
Eurographics Conference on Visualization (EuroVis) 2012
S. Bruckner, S. Miksch, and H. Pfister
(Guest Editors)

Volume 31 (2012), Number 3

Employing 2D Projections for Fast Visual Exploration of Large Fiber Tracking Data

Jorge Poco[1,2], Danilo M. Eler[1,3], Fernando V. Paulovich[1] and Rosane Minghim[1]

[1]ICMC/USP, São Carlos/SP, Brazil
[2]Polytechnic Institute of New York University, USA
[3]FCT/Unesp, Presidente Prudente/SP, Brazil

Abstract

Fiber tracts detection is an increasingly common technology for diagnosis and also understanding of brain function. Although tools for tracing and presenting brain fibers are advanced, it is still difficult for physicians or students to explore the dataset in 3D due to their intricate topology. In this work we present a visual exploration approach for fiber tracts data aimed at supporting exploration of such data. The work employs a local, precise and fast 2D multidimensional projection technique that allows a large number of fibers to be handled simultaneously and to select groups of bundled fibers for further exploration. In this approach, a DTI feature dataset, including curvature as well as spatial features, is projected on a 2D or 3D view. By handling groups formed in this view, exploration is linked to corresponding brain fibers in object space. The link exists in both directions and fibers selected in object space are also mapped to feature space. Our approach also allows users to modify the projection, controlling and improving, if necessary, the definition of groups of fibers for small and large datasets, due to the local nature of the projection. Compared to other related work, the method presented here is faster for creating visual representations, making it possible to explore complete sets of fibers tracts up to 250K fibers, which was not possible previously. Additionally, the ability to change configuration of the feature space representation adds a high degree of flexibility to the process.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques— I.3.8 [Computing Methodologies]: Computer Graphics—Applications H.5.0 [Information Interfaces and Presentation]: General—

1. Introduction

Diffusion Tensor Imaging (DTI) quantifies, using a tensor field, the diffusion of the movement of molecules in biological tissues caused by intern thermal energy. Some tissues restrict this movement decreasing molecule displacement. Tissues that allow molecules to travel easily in a particular direction are called anisotropic. An example is the brain white matter. In the white matter the water molecules move faster in the direction of the length of the neuron cells, composing the brain fibers. DTI has gained importance in the last years due its capability to reflect direction of neuron fiber bundles and its relation to brain functionality.

Techniques employed to visualize DTI datasets can be classified in 3 main categories. In the first, glyphs are used to represent tensors. Some glyph visualizations generate good visual representations, but interpretation of the underlying fiber structures is limited due to clutter in visual space.

Tensor data can also be converted to scalars, and in this case classical volume visualization techniques are employed. Using scalar visualization it is possible to find areas of interest, such as a tumor inside the brain. The disadvantage is that we loose information when reducing the tensors to one scalar, such as directionality.

In the third category of visualization for DTI, tensors are converted into vectors using spectral decomposition. Tracing streamlines using these vectors correspond to finding directions of the brain fibers, the fiber tracts, which are three-dimensional pathways representing characteristics of water diffusion. Advantages of this group are the global view of the diffusion inside the brain and of the connectivity between different brain zones. This visualization is widely used in the analysis of DTI datasets. However, exploration in visual space remains a challenging task due to the large number of derived pathways and the complexity of fiber space geometry [CDZ*09].

Information visualization techniques have been proposed to relieve analytic problems of multi-valued volume data with complex topology through mappings associated with feature spaces. Specifically to the case of fiber tracts visualization, Chen *et al.* [CDZ*09] and Jianu *et al.* [JDL09] have used multidimensional projection techniques as an effective way to explore the underlying feature space and find groups of similar fibers as well as the relationships between such groups. In this work we also adopt multidimensional projection techniques to explore DTI datasets, but refine the approach by improving the layout of fiber features and proposing a novel interactive approach were the user can explore large fiber tracts using samples and rearrange the layout as his or her analysis progresses. In addition, a new visual representation of fiber bundles using surfaces is suggested, offering an alternative view of bundle structures. The main contributions of our work are listed below:

- a process to explore large collections of fiber tracts employing a fast and precise multidimensional projection technique of the fibers feature space;
- an exploration strategy whereby the user is in control of the final outcome of the projection. With that strategy the entire fiber dataset can be explored at once starting from a sampled fraction of it;
- a new and fast approach for generating surface representations of fibers sets, aiming at improving the perception of bundles in object space when compared to other common representations, such as lines or tubes;
- evidence that curvature features improve clustering of fibers in feature space.

In Section 2 we discuss related work and give evidence of our contributions. Section 3 describes the methods used for fiber feature extraction and the multidimensional projection technique responsible for creating the visual representation. The attained results are presented and discussed in Section 4. Finally, in Section 5 we draw conclusions and discuss future work.

2. Related Work

The interactive analysis of volumetric datasets involves very challenging tasks for multivariate volume visualization, where each voxel is described by more than one scalar, vector or tensor. Due to the ability of information visualization techniques to explore multidimensional information, efforts are being made to integrate them in the analytical processes that were usually addressed by scientific visualization methods.

In this section we describe contributions with the fruitful purpose of combining techniques from the information visualization field into a volumetric visualization environment, an increasingly common trend.

One of the first proposed solutions in this area is the WEAVE system [GRW*00]. WEAVE provides an explicit coordination between volumetric visualization and bidimensional scatterplots via linked-brushing. The SimVis system [DGH03] was developed based on the same ideas, with the addition of a language to formalize subsets that can be instantiated by the linked-brushing mechanism.

More recent techniques try to handle the complete feature space at once. Akiba and Ma [AM07] suggest a three-spacial visualization joining parallel coordinates [Ins85] representing the feature space, with volumetric rendering in the object space, coupled with uni-dimensional plots showing the evolution in time of a selected property. Similarly, Blaas *et al.* [BP08] define a fast approach to render massive parallel coordinates, supporting analysis of larger scientific datasets in real time.

Linsen *et al.* [LVRR08] have proposed an approach based on surface extraction from multidimensional volume data. The surface segments the data according to a multi-variate function over the feature space. In this approach, the feature space exploration is done using a hierarchical clustering method based on density, whose clusters are shown as density level sets in a layout constructed by the *3D star coordinates* technique [SY06]. Upon this layout, the user can select the clusters that correspond to surfaces of interest in object space.

In another work, Linsen *et al.* [LVLR09] present a visual approach based on surface extraction from multi-variate volumes. The guideline for the surface extraction is decided based on an analysis of the feature space using a hierarchical cluster, shown as a tree in a 2D radial layout. In this representation, the user can select groups of interest and observe their features using a parallel coordination representation.

We apply this combination of multidimensional visualization and scientific visualization methods for the purpose of supporting interaction with data sets consisting of brain fiber tracts. For the analysis of fiber tracts, an interactive method was introduced in [CDZ*09], which improves the exploration of the 3D fibers representation using a 2D representation containing a low-dimensional embedding of the DTI fibers. The 2D representation is created using a multidimensional scaling technique, preserving the spatial relationships of the fiber tracts and providing an uncluttered representation of the data. This reduces the time spent on the analysis and the mental workload in recognizing 3D DTI fibers. In [JDL09] a similar visual exploration paradigm was presented that facilitates navigation through fiber tracts by combining a traditional 3D model viewing with two information visualization representations. One of them shows a hierarchical cluster tree of the fibers, and the other one is a projection of the data using a multidimensional scaling technique.

Regarding of the benefits of projecting fibers and bundles as resource to handle complexity, preforming selection in object space remains an issue. An extension of the selection techniques of the above mentioned systems was recently presented, which employs boxes interact with the 3D objects. In that work [CCA*11], multiple blocks are combined with boolean operations and widgets to select fibers. The benefits are that the user can manipulate the fibers and refine their selection in the 3D view. Boxed shape probes have the ability to be precise in terms of subspace location of selected individuals, but interacting with whole fibers or fiber bundles with geometries that sometimes diverge in object space remains challenging.

Another contribution to the problem of interacting with fiber tracks was recently presented by Demiralp *et at.* [JDL11]. Their work employs 2D paths to represent bundle of fibers in a 2D neuronal map. First the whole set of fibers is clustered; for each cluster a stylized 2D line is created that is the abstraction of this cluster. These 2 maps are created for the 3 axes, and the user employs them to select fibers that are synchronized with the tube representation in the 3D view. The advantage of this approach is that part of the anatomical information is preserved in the projection, which makes the exploration more intuitive for a new user. The problem with the approach is to split the cluster when and if necessary, since clustering is not always precise regarding fiber tracts and their association with brain function.

In the literature there is also much work aiming at automatically cluster fibers, but there is still the necessity to create new ways to display and explore these clusters. In the last years a few approaches were proposed to tackle this problem. For example, Chen *et al.* [CZCE08] employ surfaces enclosing clusters of fibers resulting from a hierarchical clustering algorithm. These surfaces are created using the alpha-shape algorithm considering the set of points from each fiber. Additionally, they include some principal fibers inside the surface. Latter Goldau *et al.* [GWH*11] proposed a technique to visualize tensor-derived parameters, such as fractional anisotropic. This is done bounding the fibers with a surface and using a color-coded slice that moves in the directions of the fibers inside the surface. The color used in the slice represents the tensor-derived parameter. Otten *et al.* [OVvdW10] proposed another way to visualize a group of fibers as a whole. They use illustration techniques to reduce geometric complexity and emphasize high-level structures. In that work, fiber bundles are represented using the silhouette, contours and hint lines representing the fibers, but this technique lacks depth perception.

The approach proposed here is similar to two of previous ones [JDL09, CDZ*09] in some aspects. We also combine projections of feature spaces with object space visualizations. Our approach adds to the available resources in a number of other aspects. It can quickly handle datasets that are one to two orders of magnitude larger. We do that by employing a recently published and faster projection technique named *Local Affine Multidimensional Projection (LAMP)* [JCC*11]. In addition, we support the analysis based on samples, placing the user in the loop by influencing and improving the projection layout, thus minimizing the number of iterations between projecting and brushing during data analysis. LAMP is also capable of distinguishing groups of bundled fibers with high precision. In contrast to other local multidimensional projection techniques, which also make use of control points as a starting point, LAMP can precisely map instances to the visual space using fewer control points. It speeds up user manipulation and organization of the sample projection employed to generate the final layout.

Apart from the advantages in scalability and precision, our system implements different visual representations for a group of selected fibers. For example, we make use of fibers' color to represent a group. We can also represent the fiber using tubes, which give us a better depth perception of the spatial position of each fiber, although this representation increases geometry complexity reducing interactivity. The third way to represent our selected fibers is surface extraction, enhancing the ability to recognize fiber tracts with potential for reducing even more the effort imposed on the user in recognizing existing bundles.

In the following sections we describe the methods employed in this work for performing feature extraction and multidimensional projection, as well as the interaction approach.

3. Improved projection based approach for the analysis of fiber tracts

In order to create the visual representations using a projection technique, it is necessary to calculate dissimilarities amongst the objects under analysis. To accomplish that for fibers, we convert them into feature vectors; then the dissimilarities between fibers are calculated using Euclidean distance between the vectors. In this work we extract features based on spatial and curvature information.

The spatial features, f_{spt}, comprise the coordinates of the start and end points of each fiber, its center of mass and the fiber length. The curvature features, f_{crv}, are calculated using the *Discrete Fourier Transform (DFT)*. First, a few points are spatially sampled over a fiber pathway, then a 1D DFT is applied on the components (coordinates) x_i, y_i, z_i of each point; where i varies from 1 to k, and k is the number of sampled points on the fibers. It results in three vectors of size k with the spectral coefficients of each component. We only use the magnitude (real part) of the coefficients. From those vectors we get the fibers high-frequency values since they contain most of the information about finer details of the fibers shape [PGD*05]. In our examples, the resulting features are 30-dimensional vectors, 10 from each component.

The final set of features is obtained by joining both sets of features $f = \alpha \cdot \{f_{spt}\} \cup \beta \cdot \{f_{crv}\}$. The α and β parameters are weights for the final features according to the importance of the information conveyed by each set of features. Spatial features code information about the spatial location and proximity between fibers, and curvatures features are useful for describing their pathways. Experimentally, we have discovered that the best values are $\alpha = [3,5]$ and $\beta = 1$.

To project these feature vectors on the visual plane we employ the *Local Affine Multidimensional Projection (LAMP)* [JCC*11] technique. We have tested several different projection approaches (see Section 4.5), and LAMP presents the best trade-off between precision and computational performance. Considering that the m-dimensional features are embedded into a $\mathbb{R}^m$ space, LAMP computes a set of representative samples, the *control points*, which are projected from $\mathbb{R}^m$ to the visual space with a fast and precise multidimensional projection technique. The final projection is attained building affine mappings, one for each instance to be projected, based on the position of the control points. The flexibility of LAMP resides in the fact that its formulation take into account only control points in a neighborhood of each instance to be projected. That is an interesting property, allowing the user to control the final projection via manipulation of a small portion of the entire dataset (the control points). In this work, representative Fibers can be manually positioned in the visual feature space, affecting the separation of fiber bundles in final projection.

Next, we present the application of this approach, starting with a concise analysis of the resulting feature space, and finishing with the exploration of a large fiber dataset.

4. Results

The results presented in this section were generated using 3 different fibers datasets. Table 1 details their content.

4.1. Analysis of Feature Spaces

The first issue when projecting a fiber dataset under analysis is how to identify the best set of features to describe that dataset. To that purpose, our framework is capable of aiding comparison of distinct feature spaces. To guide the decision on which feature set better groups fibers we employ the cluster silhouette coefficient [TSK05], which measures both cohesion and separation between clusters, over projections of different feature sets of pre-classified fibers. The best silhouette coefficient indicates the most discriminant feature space [JCC*11]. In this work, the silhouette of a projection is given by the average of the silhouettes of all instances. With this methodology we have chosen the types of fiber features to use.

Table 1: *Fibers datasets employed in the tests.*

Name	Description
PBC	dataset obtained from the 2009 Pittsburgh Brain Competition (PBC) Brain Connectivity Challenge (http://pbc.lrdc.pitt.edu/). It is composed by 250,000 fibers, 19,000 of which classified into 8 different classes.
CHEN	a human brain dataset composed by 1,248 fibers, provided by [CDZ*09]
JIANU	a human brain dataset composed by 690 fibers, provided by [JDL09]

(a) Spatial feature space.

(b) Curvature feature space.

(c) Combined feature spaces.

(d) Fibers Visualization.

Figure 1: *LAMP projections from distinct feature spaces. Their silhouette coefficients are* 0.5054 *(a),* 0.5482 *(b) and* 0.5494 *(c), indicating that the combination of spatial and curvature features render the best discrimination amongst the groups of classified fibers.*

Figure 1 shows LAMP projections for three different feature sets extracted from the 19,000 classified fibers of the **PBC** dataset. We use 50% of the nearest control points of each projected instance to build the projections. Figure 1(a) shows the result using only spatial features. Figure 1(b) shows the result using only curvature features, and Figure 1(c) shows the result of combining both sets of features. The silhouette coefficients the projections are 0.5054, 0.5482 and 0.5494 respectively, matching the visual inspection. Figure 1(d) shows the fiber dataset visualization with the same colors employed in the projections.

(a) Lines (b) Tubes (c) Surfaces

Figure 2: *Comparison of fiber visualizations using conventional lines, tubes, and our surface approach. The depth information provided by surfaces is a helpful hint to identification of intricate fiber bundle geometries.*

4.2. Enhancing Fiber Visualizations

After projecting fibers feature spaces, the proposed exploratory approach is based on linked views (see Figure 3), in which the fibers in object space are highlighted according to a selection of points on the projection. Conventionally, fibers in object space are drawn using lines or tubes. However, with such visual representation it is often difficult to see all the pathways defined by a group of fibers due to interference with other groups from certain viewpoints. Although tubes resolve the problem of depth perception of lines, they also increase the geometric complexity of the model, decreasing the level of interactivity as the number of selected fibers increases. To overcome both the geometric complexity of models and the speed of interaction, we visualize groups of fibers based on surfaces. Figure 2 shows a comparison between distinct visual representations for the **JIANU** dataset. Surface representations can convey depth information better than tubes or lines, easing recognition of fibers pathways without increasing the geometric complexity. In all pictures in this paper, objects in gray are fibers that are not highlighted, and drawn as lines.

The method to compute a fiber bundle surface is summarized in Algorithm 1. This is a sped-up version of the approach proposed in [LVRR08]. This algorithm is divided in two steps. First, we need to have a continuous representation of the group. This continuous field is created using a radial kernel around each point that forms the fibers. Linsen *et al.* [LVRR08] propose to compute the Δ (voxel size) using a minimal spanning tree to guarantee the connectivity in the extracted surface, but this procedure is very expensive, especially in this kind of problem, where we have hundreds or thousands of fibers composed by hundreds of points each. In this case, it is not essential to guarantee full connectivity since having and outlier fiber rendered using a different surface is not critical and, in fact, may help recognize important topological changes. Thus, we employ a static grid with a predefined size that embeds every point from the fibers. In our tests we have found that a grid with resolution of 100 will suffice. Following that gridding procedure, we generate an isosurface with an isovalue c. Empirically, we have found that a good value for c is $\frac{9}{16}$.

Algorithm 1 Surface creation from a fiber set.

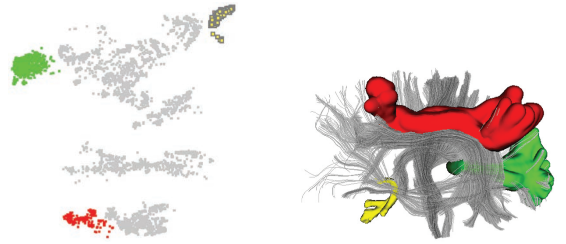
Require: F : selected fibers.
Ensure: $Surf$: a mesh of triangles.
procedure CREATESURFACE(F)
  $P = \{Points \in f_i, \forall f_i \in F\}$
  $Grid \leftarrow$ GENERATEGRID(P).
  $Surf \leftarrow$ MARCHINGCUBES($Grid, c$).
end procedure

Require: P : points.
Ensure: $Grid$: a grid embedding the points.
procedure GENERATEGRID($Points$)
  $dim \leftarrow 100$
  $[p_{min},\ p_{max}] =$ BOUNDINGBOX($Points$)
  $\Delta \leftarrow \frac{p_{max} - p_{min}}{dim}$
  $Grid \leftarrow$ Volume of size (dim, dim, dim)
  for all $p_i \in P$ **do**
    $C \leftarrow \frac{p_i - p_{min}}{\Delta}$ ▷ Sphere center.
    $R \leftarrow 5$ ▷ Sphere radius.
    for all $v_j \in$ SPHERE(C, R) **do**
      $d \leftarrow$ DIST(v_j, C) ▷ Euclidean distance.
      $v_j \leftarrow v_j + (\frac{1-d}{R})^2$
    end for
  end for
end procedure

4.3. Linked views

One important feature of the interaction in this context is linking a projection of the feature space to the 3D object spacial view of the fibers. This linked-brushing mechanism aims at verifying if groups on the projected space represent fibers with similar shapes and shared regions of object space. This process is illustrated in Figure 3, where three groups of fibers are selected and colored in the projection. The fibers belonging to these selected groups are highlighted in object space with the same colors.

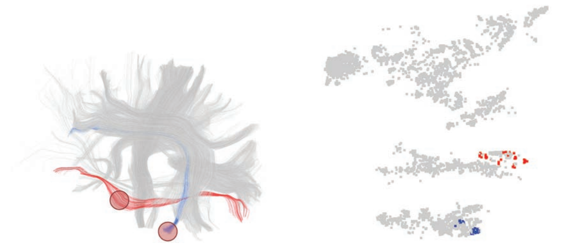
The coordination in the opposite direction is available too. Under selection of fibers in 3D object space, one can find fibers with similar features in the feature space projection. However, due to the large amount and the overlapping of fibers in different bundles, it is not possible to select all the similar ones in the object space. In this case, as shown in Figure 4, small groups or individual fibers can be selected in object space (Figure 4(a)), with their corresponding points highlighted in the projected space (Figure 4(b)). In this view,



(a) Selection of fibers in the projected feature space.

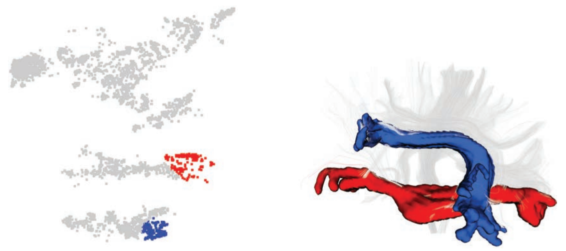
(b) Highlight in object space of fibers selected in (a).

Figure 3: *Coordination from projected 2D space (a) to object 3D space (b).*



(a) Selection in object space.

(b) Selected fibers in projected space.



(c) Selection in the neighborhoods of the highlighted points in (b).

(d) Highlight in object space of the fibers selected in (c).

Figure 4: *Two-way coordination between the object and projected spaces to guide the selection of similar fibers.*

one can use these highlighted points as a guide to select additional neighboring points (Figure 4(c)). The fibers represented by these selected neighborhoods are then highlighted in object space (Figure 4(d)), showing additional fibers, similar to the ones originally selected (see Figure 4(a)). This functionality reinforces the advantages of using multidimensional projections to give users access to fibers based on similarity of features.

4.4. Exploration of Larger datasets

At the beginning of the exploratory process there may be hundreds of thousands of points and fibers. Even with ex-

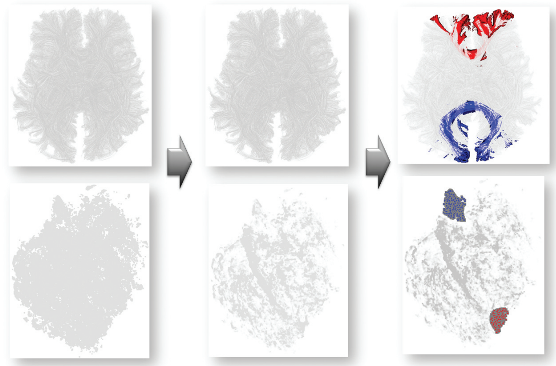


Figure 5: *Example of an exploratory scenario of a large brain fiber dataset. Transparency can be used to enhance the ability to locate dense regions on projections.*

tremely fast projections, if we inspect the **PBC** 250,000 fibers dataset shown in Figure 5 (left), there is large overlapping of points. In this section we present two approaches to explore datasets such as these.

4.4.1. Use of transparency

To enable better exploration in overloaded views one strategy is to apply transparency to the points on the projection allowing the identification of density of a region. The result of this process is shown in Figure 5 (middle views). Denser regions are more opaque and can direct the start of the exploration of large datasets. In Figure 5 (right views), we select two distinct groups of points and assign different colors to them, coloring the corresponding fibers accordingly.

Alternatively, one can apply a cluster algorithm to the projection and represent each cluster of fibers with a different color in both views. This strategy supports multilevel exploratory tasks starting with an overview until finding small clusters of interest.

4.4.2. Sample-based exploration

To relieve the cognitive overload involved in exploring too many data points at once, our approach allows the user to influence the final layout of the projection by interacting with a small sample of projected fibers first.

We perform this by selecting a small sample of fibers from the complete dataset, and projecting them to the visual space. In the examples here the samples are randomly selected. These samples can be explored and their position on feature space changed, driving the projection of the remaining points. Figure 6(a) shows the projection of a small sample with 500 instances of the **PBC** feature dataset. A selected group of points and the corresponding fiber pathways are colored in blue.

Since LAMP allows users to interfere with the projection process by manipulating small samples (the control points), exploration of a large dataset can be sped-up and more easily attained. For instance, Figure 6(b) presents two groups separated in the sample projection, the green and red groups, that may be considered similar when the object space is inspected. If the user wishes, he or she can manually join both groups into one group. Figure 6(c) shows the userst' change, colored in red and highlighted in object space. This initial manipulation is then propagated to the final projection, since the geometry of the final projection is always guided by the projection of the control points. This type of control allows the user to correct imprecisions caused by preprocessing steps without data handling.

In addition, the initial color assigned on the sample projection can be propagated to the final projection. Here we accomplish this by simply assigning, to each new projected instance, the color of the most similar control point. Figure 6(d) presents the final projection of the complete **PBC** dataset. This initial coloring can then guide the exploration of the full dataset.

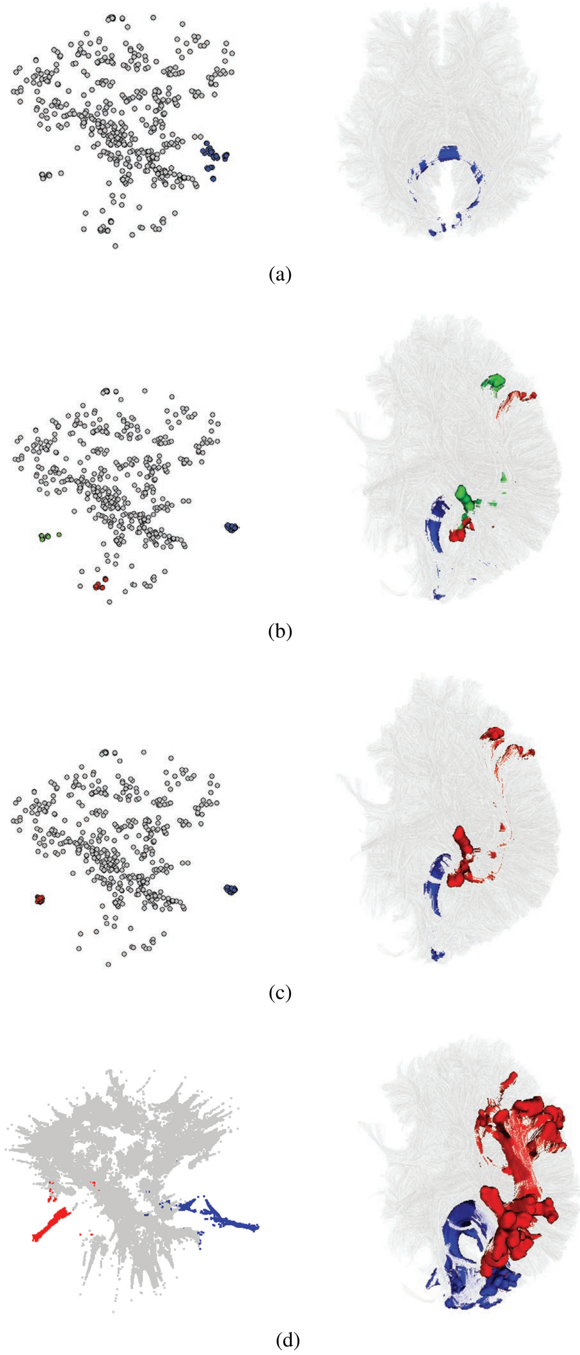
In the previous example, from Figure 6(b) to Figure 6(c), the user joined the green and red groups. Alternatively, if the user considers that these two groups belong to distinct classes of fibers, he or she could manipulate the control points as shown in Figure 7(a). The final projection generated from this manipulation is shown in Figure 7(b) and the fibers highlighted with the same color used in projection are shown in Figure 7(c). This example shows the robustness of LAMP to preserve changes expressed by user, even when separating similar fibers (e.g., fibers belonging to the same class). This is feasible by means of the truly local formulation of LAMP, which takes into account only control points in a neighborhood of each instance for the mapping. In these two examples, only 25% of the nearest control points of each projected instance is employed to build the final mappings.

4.5. Comparison with previous approaches

One of the main differences between our work and previous ones is that we have adapted the process to accommodate much larger datasets and have improved interactivity by employing LAMP. This improvement in scalability and interactivity is attained with no loss in precision measured by stress or by the silhouette coefficient.

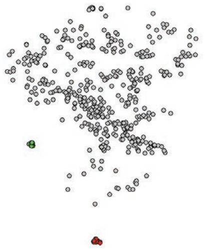
Figure 8 presents a projection of the **CHEN** dataset using our approach, as well as the one produced by the approach defined in [CDZ*09], with the fibers highlighted in the same colors as their selected groups.

Another way to evaluate projections is by considering how they reflect the given feature space. Figure 9(a) shows the distance plot produced by LAMP. It plots the distances in original feature space against distances in projected space. It

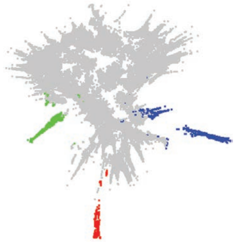


(a) (b) (c) (d)

Figure 6: *Exploring a large fiber dataset using a sampling strategy via LAMP. In (a) and (b) The user starts by projecting and exploring a small portion of the dataset. Groups can be set apart or joined together (b). Then the rest of the dataset is projected and colored using these changes as guidance (d).*



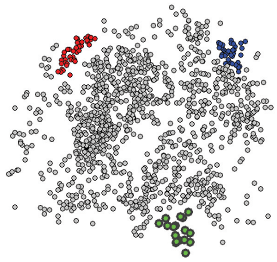
(a)



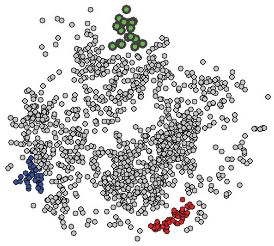
(b)

(c)

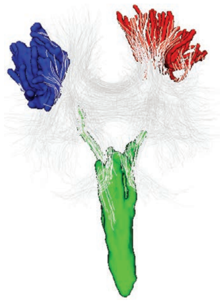
Figure 7: *An alternative sample projection manipulation for exploration shown in Figure 6. Even though the samples in red and green groups belong to the same class, LAMP is capable of preserving the layout established by the user.*



(a) 2D projection using LAMP.

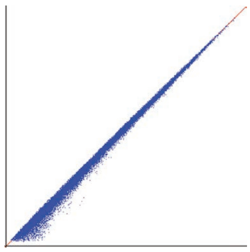


(b) 2D projection from MDS technique used in [CDZ*09].

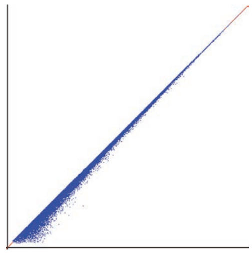


(c) Fiber visualization of the **CHEN** dataset.

Figure 8: *Projections created using our approach (a) and the method in [CDZ*09] (b), yielding similar point placements with highly differing processing times. (c) shows the highlighted fibers with the same colors employed in the projections.*



(a) Distance plot for LAMP.



(b) Distance plot for [CDZ*09].

Figure 9: *Distance plots for our approach (a) and [CDZ*09] (b), both based on our feature space.*

can be seen that the LAMP projection reconstitutes the original distances quite well, that is, if the chosen features are capable of describing groups in the dataset, so will the projection. Figure 9(b) presents the distance plot when the projection technique is the same of the employed on the Chen's work, the classical scaling (MDS) technique. Since we do not have the features of Chen's work, as an analysis exercise, we create an MDS projection using our features just to compare the precision of LAMP with the MDS approach.

The difference in precision between Classical MDS and LAMP is not significant, but MDS is computationally much more expensive. In fact, in this case LAMP is indeed more precise. The MDS projection of Figure 8 presents 0.00112 of stress against 0.00102 attained by LAMP. Table 2 presents the running times obtained on an Intel® Core™ i7 CPU 920 2.66GHz, with 8GB of RAM, comparing LAMP with other state of art techniques, that is, Piecewise Laplacian-based Projection (PLP) [PEP*11] and Part-Linear Multidimensional Projection (PLMP) [PSN10], and the MDS technique used on Chen's [CDZ*09] work. All of them are implemented in Java. We were not able to run the full **PBC** dataset with the MDS technique on the mentioned computer. It is possible to notice that LAMP is some orders of magnitude faster than MDS. It is slower than PLMP, but it is much more flexible to handle the final projection through the control points due to its local nature. Furthermore, LAMP requires less control points than PLP and PLMP. This is an important aspect of our user centered approach, since better control over the final projection can be attained by handling less information. We refer to [JCC*11] for comparison between LAMP and other projections.

To compare visual quality between projections in our case, we compute the silhouette coefficient for PLP, PLMP, MDS and LAMP, as shown in Table 3. We have employed

Table 2: *Running times (in seconds) comparing our approach and other approaches.*

Dataset	PLP	PLMP	LAMP	MDS
PBC (250K)	273.07s	1.01s	7.26s	—
PBC (19K)	1.01s	0.06s	0.24s	105.33s
JIANU	0.11s	0.02s	0.02s	0.13s
CHEN	0.04s	0.02s	0.01s	0.32s

the labeled **PBC** dataset with 19,000 fibers, considering three different feature spaces: curvature, spatial, and both combined. There is no significant difference between the measured values, indicating that amongst those techniques the capability of locating groups of fibers is similar. The improvement of the curvature feature over the spatial feature is, however, clear in some cases.

Table 3: *Comparison of the group quality using silhouette coefficient of the techniques shown in Table 2.*

Features	PLP	PLMP	LAMP	MDS
Curvature	0.6050	0.4826	0.5482	0.5354
Spatial	0.5496	0.5025	0.5054	0.5251
Both Combined	0.6040	0.6269	0.5494	0.5359

We could not compare to the method proposed in [JDL09] because we did not have access to their proposed tool.

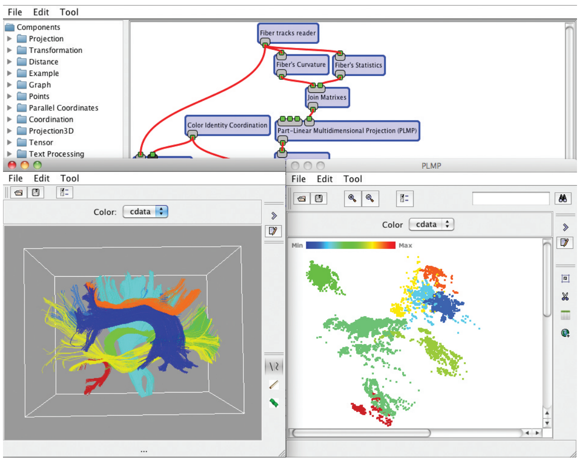


Figure 10: *Main window of the developed system to explore fiber tracts datasets.*

4.6. The System

All the needed algorithms and methods were implemented as components inside our visualization framework, which allows users to create data transformation pipelines according to their needs. The core of this system is composed by data readers, data writers, multidimensional projections, and views in 2D and 3D. In our case, we have developed extra components to manipulate fiber data, which extract feature vectors, create surface views of fiber bundles and visualize fibers pathways in object space. All these views are linked, and interaction in one view propagates to all others.

Figure 10 shows the main screen of the framework with the pipeline used to create two views for a fiber tracking dataset. In one view we present the classical representations as lines in 3D and colored using class information. In the second view we can see the 2D projection.

5. Conclusions

We have presented a strategy for the exploration of Diffusion Tensor Imaging datasets, as well as a visualization framework made freely available. We have focused in exploration based on fiber tracts.

In this work we employed spatial and curvature features to define feature spaces, and LAMP projection technique to represent the feature space in a projected space. Compared to previous strategies, our employment of LAMP as a projection tool has made the interaction with very large fiber datasets possible, with capability of exploring 250,000 fibers, with visual quality equal or superior when compared to previous approaches. We have also shown evidence that curvature features improve visualizations compared to spatial features alone.

Additionally, LAMP adjusts very well to positioning, in visual space, of sampled points. This capability allows the user to organize groups of fibers based on projected feature space and relate that to his or her findings, providing, therefore, a novel, fast, and intuitive way for the user to control possible limitations of the projection process and of the feature space. Users can adjust positioning of groups of fibers in the feature space to match their position in object space and propagate those changes to a projection for the full dataset.

The improvement in scalability and flexibility provided by this work validates projections as a promising tool to explore complex 3D spaces via 2D feature space representations. For DTI data this is particularly promising, but the approach can also be used in any volumetric real or simulated data from vector or tensor data, in need for new forms of access to volumetric object space.

Acknowledgments

The authors acknowledge the financial support of the Brazilian financial agencies CNPq and FAPESP.

References

[AM07] AKIBA H., MA K.-L.: A tri-space visualization interface for analyzing time-varying multivariate volume data. In *EuroVis – Eurographics / IEEE VGTC Symposium on Visualization* (2007), Eurographics Association, pp. 115–122. 2

[BP08] BLAAS J., POST C. P. B. F.: Extensions of parallel coordinates for interactive exploration of large multi-timepoint data sets. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1436–1451. 2

[CCA*11] CAI H., CHEN J., AUCHUS A. P., CORREIA S., LAIDLAW D. H.: Inbox: In-situ multiple-selection and multiple-view exploration of diffusion tensor mri visualization. Abstract at: 1st ieee symposium on biological data visualization, 2011. 3

[CDZ*09] CHEN W., DING Z., ZHANG S., MACKAY-BRANDT A., CORREIA S., QU H., CROW J., TATE D., YAN Z., PENG Q.: A Novel Interface for Interactive Exploration of DTI Fibers. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1433–1440. 2, 3, 4, 7, 8

[CZCE08] CHEN W., ZHANG S., CORREIA S., EBERT D. S.: Abstractive representation and exploration of hierarchically clustered diffusion tensor fiber tracts. *Computer Graphics Forum 27*, 3 (2008), 1071–1078. 3

[DGH03] DOLEISCH H., GASSER M., HAUSER H.: Interactive feature specification for focus+context visualization of complex simulation data. In *VISSYM '03: Proceedings of the symposium on Data visualisation 2003* (Aire-la-Ville, Switzerland, Switzerland, 2003), Eurographics Association, pp. 239–248. 2

[GRW*00] GRESH D. L., ROGOWITZ B. E., WINSLOW R. L., SCOLLAN D. F., YUNG C. K.: Weave: A system for visually linking 3-d and statistical visualizations applied to cardiac simulation and measurement data. In *Proceedings of the 11th IEEE Visualization Conference (VIS 2000)* (Washington, DC, USA, 2000), IEEE Computer Society. 2

[GWH*11] GOLDAU M., WIEBEL A., HLAWITSCHKA M., SCHEUERMANN G., TITTGEMEYER M.: Visualizing dti parameters on boundary surfaces of white matter fiber bundles. In *Proceedings of the Twelfth International Association of Science and Technology for Development International Conference on Computer Graphics and Imaging* (February 2011), Zhang J., (Ed.), ACTA Press, pp. 53–61. 3

[Ins85] INSELBERG A.: The plane with parallel coordinates. *The Visual Computer 1* (1985), 69–91. 2

[JCC*11] JOIA P., COIMBRA D., CUMINATO J. A., PAULOVICH F. V., NONATO L. G.: Local affine multidimensional projection. *IEEE Transactions on Visualization and Computer Graphics 17* (Dec. 2011), 2563–2571. 3, 4, 8

[JDL09] JIANU R., DEMIRALP C., LAIDLAW D. H.: Exploring 3D DTI fiber tracts with linked 2D representations. *IEEE transactions on visualization and computer graphics 15*, 6 (2009), 1449–56. 2, 3, 4, 9

[JDL11] JIANU R., DEMIRALP C., LAIDLAW D. H.: Exploring brain connectivity with two-dimensional neural maps. *IEEE Transactions on Visualization and Computer Graphics PP*, 99 (2011), 531–532. 3

[LVLR09] LINSEN L., VAN LONG T., ROSENTHAL P.: Linking multidimensional feature space cluster visualization to multifield surface extraction. *IEEE Computer Graphics and Applications 29*, 3 (2009), 85–89. 2

[LVRR08] LINSEN L., VAN LONG T., ROSENTHAL P., ROSSWOG S.: Surface extraction from multi-field particle volume data using multi-dimensional cluster visualization. *IEEE transactions on visualization and computer graphics 14*, 6 (2008), 1483–90. 2, 5

[OVvdW10] OTTEN R., VILANOVA A., VAN DE WETERING H.: Illustrative white matter fiber bundles. *Computer Graphics Forum 29*, 3 (2010), 1013–1022. 3

[PEP*11] PAULOVICH F. V., ELER D. M., POCO J., BOTHA C. P., MINGHIM R., NONATO L. G.: Piecewise Laplacian-based Projection for Interactive Data Exploration and Organization. *Computer Graphics Forum 30*, 3 (2011), 1091–1100. 8

[PGD*05] PAZOTI M. A., GARCIA R. E., DALTON J., PESSOA C., BRUNO O. M.: Comparison of shape analysis methods for guinardia citricarpa ascospore characterization. *Electronic Journal of Biotechnology 8*, 3 (2005), 265–275. 4

[PSN10] PAULOVICH F. V., SILVA C. T., NONATO L. G.: Two-phase mapping for projecting massive data sets. *IEEE Transactions on Visualization and Computer Graphics 16* (November 2010), 1281–1290. 8

[SY06] SHAIK J. S., YEASIN M.: Visualization of high dimensional data using an automated 3D star co-ordinate system. In *International Joint Conference on Neural Networks* (2006), pp. 1339–1346. 2

[TSK05] TAN P.-N., STEINBACH M., KUMAR V.: *Introduction to Data Mining*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005. 4