

A Fiducial-Based Tangible User Interface for White Matter Tractography

Steven R. Gomez, Radu Jianu, and David H. Laidlaw

Department of Computer Science
Brown University
{steveg, jr, dhl}@cs.brown.edu

Abstract. We present a method for interacting with 3D brain tract visualizations using a webcam and a fiducial marker that can be constructed cheaply in any home or office. Our contributions are a fiducial-based tracking architecture in the context of white matter tractography, and a preliminary evaluation with domain scientists providing usability and design insights. Expert feedback indicates that model positioning in our system is easier than in previous methods using traditional input devices or two-dimensional input interfaces, and that tract selection may be faster to execute using our tool, given training and practice.

1 Introduction

Scientists can now explore the shape and connectivity of fibrous tissues, such as muscle and brain white matter, through visualizations of Diffusion Tensor Magnetic Resonance Imaging (DTI) data that commonly render variations of streamlines, such as streamtubes and hyperstreamlines, in 3D. In the case of the brain, these streamline models are visually dense as a consequence of the brain’s complex circuitry. As a result, typical interactions with white matter tracts, such as bundle selection or inspection of the model, may be difficult for a user to perform with conventional interface tools. In this paper, we present a new method for interacting with neural fiber tracts using a computer vision-based interface that allows for intuitive manipulation of the DTI model.

We use fiducial tracking to position the brain and perform 3D selection of fiber tracts. In lieu of a typical interface, e.g. keyboard and mouse, or specialized input devices, the user holds and moves a homemade marker object in front of a webcam to manipulate the model. This marker can be constructed inexpensively from a pattern and common household materials. In our experiments, we constructed a cardboard cube and decahedron whose faces are covered with paper Augmented Reality (AR) patterns, which can be printed on any printer.

Figure 1 shows a user interacting with a DTI brain model using our system. We have obtained feedback from experts in an anecdotal study for an initial prototype. Results suggest that this type of lightweight 3D interaction has the potential to enable faster interaction with dense fiber tract collections. Expert feedback indicates that our new representation is more intuitive – and may be easier to use and learn – than conventional DTI model interaction methods.

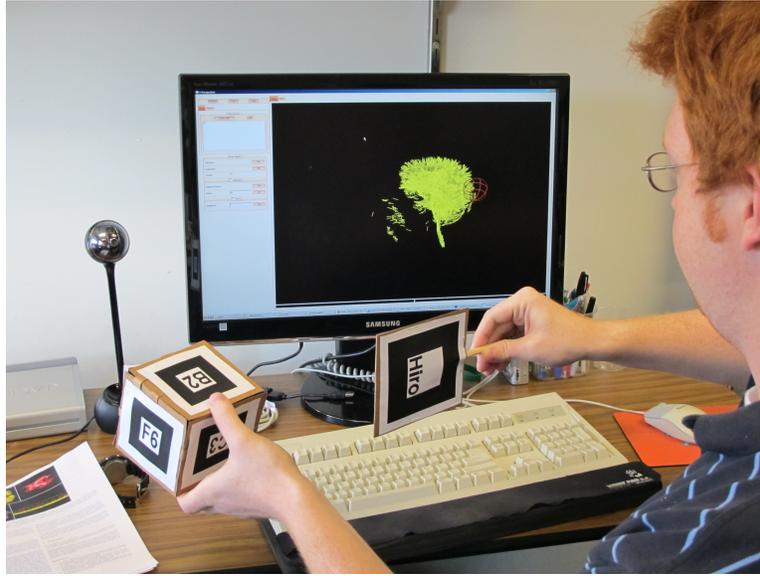


Fig. 1. The user holds both the brain marker and selection marker to identify tracts of interest in the visualization.

2 Related Work

2.1 Interacting with DTI Models

White matter tracts in the brain are frequently visualized as tubes rendered from streamline traces in DTI data [1–4]. Common interaction tasks for these visualizations include exploration of the 3D model and tract of interest (TOI) selections in the tract bundles that make up the model. Tract selections are traditionally performed directly in 3D: regions of interest (ROIs) are placed in the volume along the presumed path of the desired tract bundle; the application then selects only fibers that intersect those ROIs [5–7]. At the same time, mastering the interactions required for 3D ROI manipulation using conventional input devices, e.g. mouse and keyboard, often slows inexperienced scientists in their exploration.

Due to these limitations, recent work has explored new input techniques that allow brain scientists to perform tract bundle selection with minimal overhead. In CINCH [8], Akers presents a sketch and gesture interface for identifying tract bundles by curve similarity and other heuristics. Jianu et al. [4] and Zhang et al. [3] present methods that use tract similarity to create 2D abstractions of the DTI space along the original 3D model, and use brushing and linking to map interactions performed in 2D onto the 3D model. While simplifying interaction, these methods require scientists to operate on abstract, unfamiliar representations and maintain a mapping between the different views of the same data. Closer to our work is a technique introduced by Zhou et al. [9] for lassoing

fiber tracts into bundles using a finger-mounted haptics device that provides force feedback and 3D tracking. Recently, Klein et al. [10] created a tract selection interface that allows a user to create 3D ROIs using a Nintendo Wii remote controller. Our system, which uses fiducial marker tracking with the ARToolKit [11] and a simple webcam, differs from these by offering an interface for the brain model that allows the user to interact in 3D space, as opposed to planar sketching, but requires no specialized hardware.

2.2 Fiducial Tracking

Fiducial-based tracking has grown in popularity with the release of systems like ARToolKit and ARTag [12, 13], which provide pattern creation and recognition as library functions. Fiducials have been broadly applied in applications ranging from robot navigation and communication [14], games and learning [15], and the design of tangible user interfaces (TUIs) [16] where real objects are used as input devices. We find a novel application area for TUI design principles in interactive scientific visualizations.

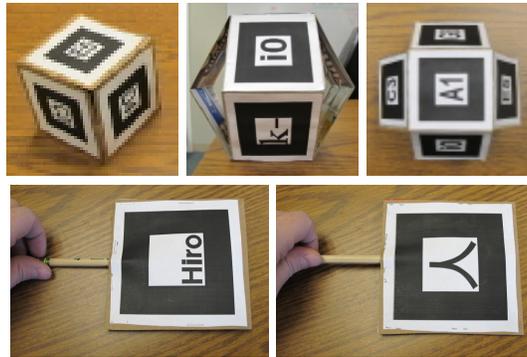


Fig. 2. Brain model markers constructed as polyhedra from cardboard and augmented reality (AR) patterns on the flat faces. Top, left to right: cube model; side view of decahedron; top view of decahedron. Bottom, faces A and B of the two-sided selection marker.

3 Methods

3.1 Brain Model Manipulation

We use the ARToolKit [11] for tracking fiducial patterns in the video feed from the webcam. Each pattern is a monochrome character sequence against a white background and inside a black square frame. We choose each sequence to be

asymmetric and unique among the other patterns to avoid as much pattern ambiguity in the detection phase as possible. The specification for each fiducial pattern we use is created from a toolkit utility and stored as a bitmap of pixel brightness intensities. During the main detection loop in our visualization software, the toolkit provides the orientation of any detected objects and a match-confidence score for each specific pattern we use. The detected object is labeled by the pattern that scores the highest confidence value.

We use the orientation of detected patterns to compute a coordinate frame that transforms the 3D model consistently with the user’s manipulation of the tangible model. For a detected pattern in the frame, ARToolKit returns orientation parameters that determine the location and surface normal of that face. By the marker’s design, we know where each pattern is oriented in object space in the prototypical marker shape. The rotations that move each specific face into a common position are pre-computed and stored in a hash table for quick lookup. As a result, we transform the model simply by composing the coordinate frame returned by the toolkit with the object space rotations stored for the respective pattern.

Because multiple patterns may be detected simultaneously from the surface of the cube or decahedron markers, the coordinate frame most similar to the existing model transformation is chosen. This similarity is determined as the arc length produced by projecting a common unit vector into both the current and detected coordinate frames. We choose the pattern that produces the minimum arc distance between these transformed points on the unit sphere. This ensures that the same detected face is used as the reference coordinate frame until it is no longer detected in the camera’s view. This continuity check is aimed at keeping the 3D model as stable as possible during manipulation. As an added measure to mitigate camera jitter, we maintain a short history of detected marker orientations (up to 10 frames) and apply an average of these coordinate frames to smoothly transform the model over time.

3.2 Selecting Tracts

The selection tool we constructed is a two-sided marker with one unique pattern on each face and one handle that can be twisted to flip the marker to the reverse sign. Detecting the selection tool is similar to detecting the brain markers, but its orientation is used for positioning an on-screen selection sphere rather than moving the brain model.

When the A-side is visible to the camera, as the user moves the selection sphere through the white matter model, tracts intersecting the sphere are highlighted as a preliminary selection. Once the user is satisfied with the selection he can twist the tool in place so that the reverse B-side is detected. Toggling this face triggers the “selection” action, which commits the preliminary selection. Figure 3 illustrates this process; preliminary selections are marked in blue, committed ones in red. Three selection modes are possible: add tracts, remove tracts and intersect. The selection mode is used to construct preliminary selections by composing the set of tracts intersecting the selection sphere at each moment in

time with the current selection. An existing selection can thus be cleaned by brushing over tracts while the “remove tract” mode is active. The “intersection” mode provides a logical AND operation for selecting tracts crossing multiple areas of interest in the model.

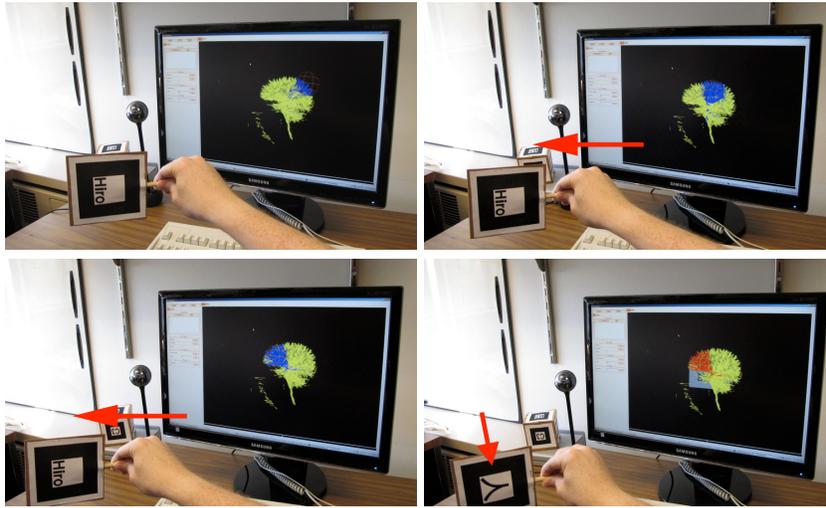


Fig. 3. Selecting tracts of interest in the brain model. The sequence shows the user positioning the selection sphere in model space by moving the physical marker, then flipping the marker to add the highlighted tracts to the selection. The marker can then be removed from the camera’s field of view and the selection persists. In the images above, the red arrows indicate movement from the previous frame.

3.3 Gesture Recognition

We designed our application to support simultaneous movement of the brain model marker and the selection marker. We hypothesized that using a mouse and keyboard for selection interactions (e.g. changing ROI radius or selection mode) might interrupt the user’s workflow, causing longer completion times for selection tasks. To overcome this limitation, we developed a set of gestures performed with the markers to execute commonly performed interactions in the system.

For flexibility, the interface also supports selection controls using traditional keyboard and mouse input, accommodating users who prefer to use only the brain model marker with the camera. When either marker leaves the field of view of the camera, interaction events are not generated by that marker. The user can effectively pause interactions by removing or obstructing a given marker, and can resume by reintroducing it into the scene.

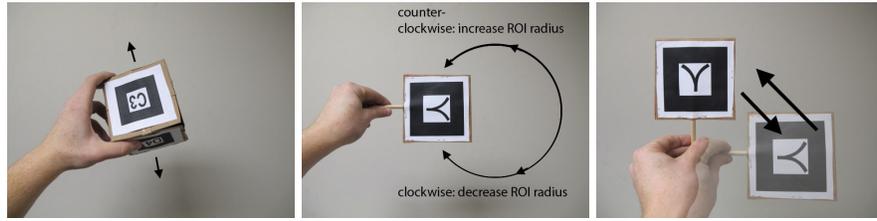


Fig. 4. Examples of gestures using the brain and selection markers. Left to right: shaking the brain marker slightly will toggle the 3D model from being frozen in position or unfrozen; creating a circular motion with the selection marker will increase (counter-clockwise) or decrease (clockwise) the radius of the selection sphere; rotating the selection marker 90 degrees and returning it back to its start position will switch selection modes between “add”, “subtract”, or “intersection”.

Committing a selection can be performed, as previous described, by flipping the selection marker to show the opposing face. Toggling between selection modes (i.e. adding to, removing from and intersecting with the current selection) is performed by rotating the marker 90 degrees about its center in the image plane and then returning it to its initial orientation. The spherical selection ROI’s radius can be increased or decreased by moving the marker along a circular path clockwise or counter-clockwise. Finally, the 3D model can be frozen in place, and unfrozen, by moving the brain marker into the desired position and shaking it slightly. These gestures are illustrated in Figures 3 and 4.

Gestures are recognized by tracking a marker’s position and orientation over short periods of time – 1-2 seconds in our implementation – and computing gesture likelihood scores for the marker orientation data collected. In the case of shaking and circular movements, the scores incorporate distance traveled by the marker, displacement of the marker, and directionality of movement. A least-squares distance error function is used to assess whether a marker’s path is circular during the gesture.

3.4 DTI Model

Our test data included a diffusion-weighted MRI volume of a normal volunteer’s brain, acquired on a 1.5T Siemens Symphony scanner with the following acquisition parameters in 12 bipolar diffusion encoding gradient directions: thickness = 1.7mm, FOV = 21.7cm x 21.7cm, TR = 7200ms, TE = 156ms, b = 1000, and NEX = 3. The DTI volume was then obtained by fitting 6 independent parameters of a single second-order tensor at each voxel to the 12 measurements from diffusion-weighted MRI volume [17].

4 Results

We evaluated our prototype anecdotally with two neuropsychologists who had experience using white matter streamtube visualizations in clinical research. Each received a demonstration and brief training on the interface and gestures before exploring the tool and selecting TOIs. We gathered feedback using a think-aloud protocol during our demonstration and the users' exploration of the tool.

The workflow that both our subjects suggested, and said they were likely to use with such an application, was to position the model in a specific pose, freeze it, and use the selection marker to perform a selection. They found that for certain types of interactions, such as model positioning using the fiducial marker, our method would be helpful in supplementing the mouse and keyboard. Both agreed that different users may have varying preferences for using the fiducials or a typical keyboard/mouse interface, for either coarse model manipulation or fine-scale TOI selection. For instance, one expert said he preferred the fiducial for coarse-scale rotation of the model, then favored switching to the mouse for precise tract selection; the other felt he could set the model position more accurately with a keyboard and mouse, but preferred the fiducial for tract selection because he could quickly select tract bundles in 3D and preview the selection.

Our subjects stated that while the selection gesture is useful, they preferred using the keyboard and mouse to alter other selection parameters (e.g. adding, removing, or intersecting modes; freezing and unfreezing the model). They noted that the selection marker was at times difficult to use with precision, given that executing the twist "select" gesture can change the marker's position, causing different tracts to be selected once the gesture is recognized. One noted that this became easier after a few minutes of practice, and that more extensive initial training may give more confidence to the user. Both agreed that adding a new mode to hide non-selected tracts would allow the user to refine the selection more easily by reducing visual clutter. Additionally, one expert suggested that other medical applications requiring quick and easy inspection of a 3D model, such as a heart or tumor visualization, could benefit from an interaction method like ours.

5 Discussion

5.1 System Robustness

One finding in our work is that a vision-based approach to model manipulation introduces challenges from external factors. Fiducial recognition relies on some level of light and camera quality beyond our control. Occlusion of the fiducial patterns can also cause recognition problems, especially when concurrently manipulating both the selection marker and brain maker. One marker may block the camera's view of the other as they try to occupy an overlapping position in model space. Even a user's grip on a marker can introduce occlusion problems. In fact, we noticed that some improved usability of the decahedron over the cube

was not due to the visibility of multiple patterns at the same time, as initially intended, but instead by the ease of holding and moving it while having at least one marker visible at all times.

5.2 Accommodating User Workflow

In the selection task, the largest challenge we identified was designing interactions that were appropriate for the user’s intended level of precision for selection. For instance, a large ROI volume that moves relatively quickly in model space when manipulated may be ideal for a coarse tract bundle selection; however, for refinements of the selection, a smaller, slower moving selection volume may be easier for the user to navigate and use without error. Ideally, we would like our system to determine the user’s tract selection goals in order to set these parameters automatically, to reduce manual specification that may slow the user down. We hypothesize that a coarseness estimate can be made from the level of model zoom in the 3D visualization; a user who is closely examining the model likely wants to make finer scale selections.

As revealed by our evaluation, the workflow suggested by the users was to position the model, freeze it, and then make a selection. We believe this is due to our subjects’ familiarity with streamtube visualizations where model positioning and selection cannot happen concurrently. We hypothesize that new users, or those with training, will prefer to perform selections while manipulating both the selection marker and the model marker concurrently. We expect this would decrease the time required to execute a selection with minimal error because positioning the brain and selection markers can be done simultaneously. This may require further extension and refinement of the gesture set to obviate all keyboard and mouse interactions.

5.3 Reaching a Broad Audience

The power of this approach lies in its affordability. Many laptops are now sold with integrated cameras, and for other personal computers, external webcams are relatively inexpensive to purchase. In the released version of our application, we will make available marker templates that can be printed, cut, and folded into cubes. Users should be able to reproduce our setup in less than half an hour from downloading our tool.

Furthermore, we plan to design meaningful pictures, such as projections of typical brain views (i.e. coronal, sagittal, axial), to distribute as fiducial patterns on the marker. Each of these images offers a preview of the model’s orientation before the marker face is presented to camera. We believe these fiducials will help new users more intuitively understand how manipulation of the marker object changes the brain visualization.

6 Conclusion

We present a lightweight, affordable interaction method for DTI brain visualizations using paper fiducial markers that are easy to create. This method allows scientists to manipulate 3D white matter models directly by manipulating a physical analogue in real space in front of a camera. Our contributions include a fiducial-based tracking architecture for interacting with white matter tractography models, and an evaluation revealing the advantages of this method and suggesting design guidelines. Furthermore, by using a physical marker that can be printed, constructed easily, and used without specialized hardware, we provide a simple method for distributing the visualization tool to a wide base of users.

References

1. Basser, P., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine* **44** (2000) 625–632
2. Mori, S., Van Zijl, P.: Fiber tracking: principles and strategies—a technical review. *NMR in Biomedicine* **15** (2002) 468–480
3. Zhang, S., Demiralp, C., Laidlaw, D.H.: Visualizing diffusion tensor mr images using streamtubes and streamsurfaces. *IEEE Transaction on Visualization and Computer Graphics* **9** (2003) 454–462
4. Jianu, R., Demiralp, C., Laidlaw, D.: Exploring 3d dti fiber tracts with linked 2d representations. *IEEE Transactions on Visualization and Computer Graphics* **15** (2009) 1449–1456
5. Catani, M., Howard, R.J., Pajevic, S., Jones, D.K.: Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage* **17** (2002) 77–94
6. Wakana, S., Jiang, H., Poetscher, N.L.M., van Zijl, P.C., Mori, S.: Fiber tract-based atlas of human white matter anatomy. *Radiology* **230** (2004) 77–87
7. Maddah, M., Mewes, A.U.J., Haker, S., Eric, W., Grimson, L., Warfield, S.K.: Automated atlas-based clustering of white. In: *MICCAI*. (2005) 188–195
8. Akers, D.: Cinch: a cooperatively designed marking interface for 3d pathway selection. In: *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, New York, NY, USA, ACM (2006) 33–42
9. Zhou, W., Correia, S., Laidlaw, D.H.: Haptics-assisted 3D lasso drawing for tracts-of-interest selection in DTI visualization. *IEEE Visualization 2008 Poster Compendium (Best Poster Nominee)* (2008)
10. Klein, J., Scholl, M., Kohn, A., Hahn, H.K.: Real-time fiber selection using the wii remote. *Proceedings of the SPIE* **7625** (2010)
11. (ARToolkit) <http://www.hitl.washington.edu/artoolkit/>.
12. Fiala, M.: Artag, a fiducial marker system using digital techniques. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on* **2** (2005) 590–596
13. Fiala, M.: Designing highly reliable fiducial markers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **32** (2010) 1317–1324
14. Dudek, G., Sattar, J., Xu, A.: A visual language for robot control and programming: A human-interface study. In: *ICRA*. (2007) 2507–2513

15. Kostandov, M., Schwertfeger, J., Jenkins, O.C., Jianu, R., Buller, M., Hartmann, D., Loper, M., Tsoli, A., Vondrak, M., Zhou, W., Fiala, M.: Robot gaming and learning using augmented reality. In: SIGGRAPH '07: ACM SIGGRAPH 2007 posters, New York, NY, USA, ACM (2007) 5
16. Kato, H., Billingham, M., Poupyrev, I., Imamoto, K., Tachibana, K.: Virtual object manipulation on a table-top ar environment. *Augmented Reality, International Symposium on* **0** (2000) 111
17. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of magnetic resonance. Series B* **103** (1994) 247-254