Camino: Diffusion MRI reconstruction and processing

PA Cook, Y Bai, M J Hall, S Nedjati-Gilani, K K Seunarine and DC Alexander

Centre for Medical Image Computing, Department of Computer Science, University College London, UK. email: camino@cs.ucl.ac.uk

Abstract. Camino is an open-source, object-oriented software package for processing diffusion MRI data. Camino implements a data processing pipeline, which allows for easy scripting and flexible integration with other software. This paper summarises the features of Camino at each stage of the pipeline from the raw data to the statistics used by clinicians and researchers. The paper also discusses the role of Camino in the paper "An Automated Approach to Connectivity-based Partitioning of Brain Structures", published at MICCAI 2005.

1 Introduction

Diffusion-weighted MRI allows us to examine the characteristics of water diffusion *in vivo*. Diffusion MRI measures the particle displacement density p in each voxel of the image volume. The anisotropy of p reflects that of the tissue microstructure. Statistics of diffusion are used extensively in clinical medicine and biomedical research.

Diffusion MRI sequences typically acquire N measurements with different "diffusion weighting" gradients in each voxel. Each measurement is a sample of the Fourier transform of p a the wavenumber $\mathbf{q}$. Several reconstruction algorithms exist to fit models of p or extract useful features (such as anisotropy or fibre orientation) from the data [1]. The most widely used model is the diffusion tensor [2], which assumes a zero-mean, Gaussian p:

$$p = G(\mathbf{r}; \mathrm{D}, t) = ((4t)^3 \det(\mathrm{D}))^{-\frac{1}{2}} \exp(-\mathbf{r}^{\mathrm{T}} \mathrm{D}^{-1} \mathbf{r} (4t)^{-1}) \tag{1}$$

where D is a symmetric, second-order tensor. The log of each measurement provides a linear constraint on D:

$$\log[A^*(\mathbf{q})/A(\mathbf{0})] = \mathbf{q}^{\mathrm{T}} \mathrm{D} \mathbf{q}\,. \tag{2}$$

Since D has six independent elements, DT-MRI requires measurements with at least six independent $\mathbf{q}$ and one at $\mathbf{q} = \mathbf{0}$, though more measurements are often made to improve the robustness of the fit.

The diffusion tensor method fails when G is a poor approximation of p. Examples of this can be seen easily in the brain when two or more bundles of fibres cross within a single voxel. Several techniques exist to reconstruct multiple fibre orientations within a voxel [1]. One method is to divide the voxel into multiple compartments such that p is a mixture of zero-mean Gaussians:

$$p = \sum_{i=1}^{M} f_i G(\mathbf{r}; \mathrm{D}_i, t)\,, \tag{3}$$

where $f_1 + f_2, \ldots, +f_M = 1$.

This paper examines the image processing steps involved in the study of diffusion-MRI and shows how the open-source Camino software suite provides the capability to extract clinically useful information from raw MRI data. In Sec. 2, we introduce the data pipeline and summarise the features of Camino. In Sec 3 we discuss the role of Camino in the companion MICCAI paper “An Automated Approach to Connectivity-based Partitioning of Brain Structures” [3].

2 Camino

Camino is an object-oriented software package for processing diffusion MR images and is written entirely in Java. Users may invoke the Java programs directly or run the wrapper (bash shell) scripts provided with the software. Advanced users can construct custom scripts or extend the source code to suit their needs.

The pipeline approach to data processing is illustrated in Fig. 1. The shell scripts realise the pipeline approach and provide a stable interface for users, which means that no knowledge of Java is required to use Camino. This section examines each stage of the pipeline in detail.

2.1 Data sources

The data source can be data from a scanner or from Camino’s data synthesiser. Data consists of N measurements in each voxel. Camino processes all data in voxel order, where the measurements for each voxel are stored together. This ordering facilitates the data pipeline model and allows each voxel to be processed independently. The data pipeline allows easy parallel processing and processing of large images with minimal memory consumption.

Data from scanners is not typically in voxel order, since the scanner acquires a measurement over the whole image volume, then acquires the second measurement, and so on for all N measurements. We call this “scanner order”. The Camino command `scanner2voxel` rearranges data into voxel order.

```
$ scanner2voxel -components 60 -numvoxels 983040 <
   SubjectA.raw > SubjectA.vo.Bfloat
```

Camino’s output is raw binary, with no header information. We therefore follow a naming convention for output, appending “.Bfloat” to the file name to identify the data format as big-endian float. The convention in Java is to read and

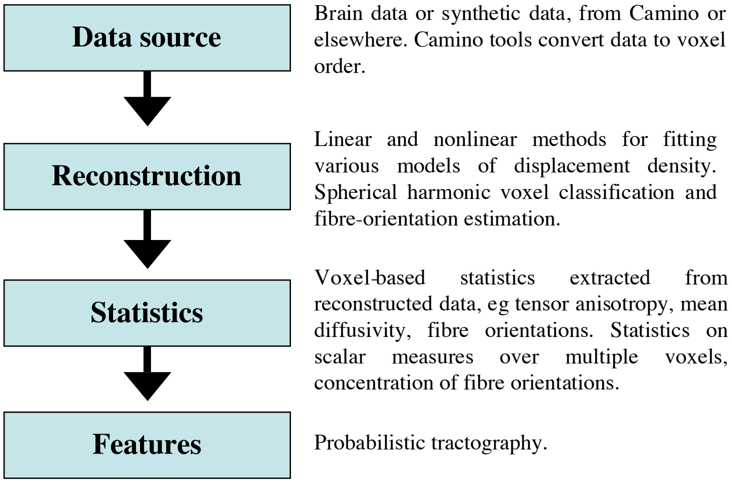


Fig. 1. The data-processing pipeline.

write big-endian data, we follow this convention in Camino. Data in little-endian format may be converted using the `shredder` program.

Every data file has an associated "scheme file", which specifies the scanner protocol used in the data acquisition. Specifically, the scheme file contains the diffusion time and wavenumber for each measurement.

Camino's data synthesizer emulates scanner sequences given a scheme file and a test function for p. Measurements can be synthesised independently for each voxel, or by bootstrap sampling from a small number of synthesised measurements. The Gaussian model (Eq. 1), and the multi-Gaussian model (Eq. 3) are implemented in Camino for use in data synthesis. Developers can add alternative models of p (eg the model proposed in [4]) by implementing the simple `ModelPDF` interface:

```
public interface ModelPDF {

    /**
     * Returns the value of the function at the specified point.
     *
     * @param x The point to sample at.
     * @return The value of the function at x.
     */
    public double at(double[] x);

    /**
     * Returns the value of the FT of the function at the
     * specified wavenumber.
```

```
     *
     * @param q The wavenumber to sample at.
     * @return The value of the FT at q.
     */
    public double ftAt(double[] q);

    /**
     * Returns a list of principal directions of the test
     * function.
     *
     * @return Array of principal directions.  The first index is the
     * principal direction number.
     */
    public double[][] getPDs();

}
```

2.2 Reconstruction

The reconstruction stage takes as input raw data and reconstructs information about p. The following command takes as input the raw data from "subject A", the associated scheme file, and fits D to the log measurements by linear regression:

```
$ dtfit SubjectA.Bfloat A.scheme > DT_A.Bdouble
```

The program outputs eight values for each voxel, in raw binary format. The first value is the error code, which is zero if the fit was successful, otherwise a nonzero code describes the nature of the error. The second value is $\log[A(\mathbf{0})]$, the estimated signal at $\mathbf{q} = \mathbf{0}$. The remaining six values are the six independent elements of D. The `dtfit` program processes a data set of $128 \times 128 \times 60$ voxels with 66 measurements per voxel in approximately two minutes on a 2.4 GHz Pentium PC.

The `dtfit` program can fit D directly to the measurements using a Levenberg-Marquardt algorithm, which offers better noise immunity [5]. Camino also supports the RESTORE method [6] for nonlinear fitting.

Camino features several high angular resolution or "multi fibre" reconstruction methods. The spherical harmonic algorithm of Alexander et al [7] detects and models non-Gaussian diffusion. The algorithm fits a nested hierarchy of spherical harmonic functions to the data and uses an F-test for deletion of variables to decide when to truncate the series. The results of the F-test identify areas of isotropic diffusion (series order 0 spherical harmonic), anisotropic Gaussian diffusion (order 2) or non-Gaussian anisotropic diffusion (order 4 or higher). The `voxelclassify` program in Camino provides voxel classification using this algorithm. Sec. 3 shows an example use of this program.

Camino implements two-Gaussian and three-Gaussian model fitting (Eq. 3). Users may optionally constrain the fitting process to improve the results. The mixing fraction f in Eq. 3 may be fixed and the fitted diffusion tensors may be constrained to cylindrical symmetry.

The persistent angular structure [8] is a special case of spherical deconvolution [9], both are implemented in the current release. These methods are computationally intensive, nonlinear inversions of the diffusion data. The Camino team is currently testing linear spherical deconvolution and the Q-ball algorithm [10] to resolve multiple fibres. Future releases will include these algorithms.

2.3 Statistics

Voxel-based statistics of the diffusion can be extracted from the reconstructed data. For tensors, two common statistics are Tr(D) and the fractional anisotropy [11]. The Camino pipeline allows users to produce these statistics from raw data by combining each stage into a single command:

```
$ scanner2voxel -components 60 -numvoxels 983040
    < SubjectA.raw | dtfit - A.scheme | fa
    -inputdatatype double > A.fa.Bdouble
```

Tr(D) is calculated in a similar manner, using the command `trd` instead of `fa`. The `dteig` program outputs the entire eigen system of the tensor. The `shredder` program extracts volumes from composite data sets. All of these programs handle one, two or three tensors per voxel. Statistics on populations of voxels include the mean and variance of scalar measures such as anisotropy, or the concentration of fibre orientations.

Camino also provides voxel and population statistics from non-tensor data, such as the output of PAS-MRI or spherical deconvolution. The `sfpeaks` program extracts peaks of spherical functions and the Hessian at those peaks for use in tractography [12].

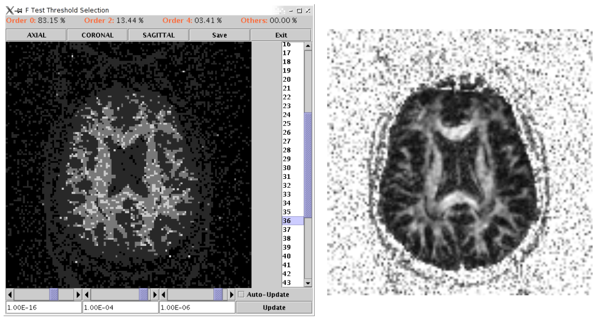
2.4 Image features

Deterministic and probabilistic tractography [13, 14] are currently being integrated into the Camino suite. These algorithms will produce streamlines or connectivity maps for connectivity analysis.

2.5 Visualisation

Camino's primary purpose is to provide a pipeline-based suite of data processing algorithms. All of Camino's output data structure is documented kept as simple as possible to facilitate integration with visualisation software. Visualisation features may be added to future releases of Camino as another stage in the pipeline, but are not a priority.

Camino currently contains only one visualisation tool, which is used to apply the F-test for spherical harmonic voxel classification. These thresholds can be determined interactively (Fig. 2) by piping the output of `voxelclassify` into the `vcthreshselect` program.


F Test Threshold Selection

Order 0: 83.15 %
Order 2: 13.44 %
Order 4: 03.41 %
Others: 00.00 %

1.00E-16
1.00E-04
1.00E-06
☐ Auto-Update

Fig. 2. The `vcthreshselect` program (left) and a slice from an anisotropy map (right).

2.6 Testing and documentation

The modular nature of Camino facilitates unit testing. Camino tests are written using the open source JUnit (http://www.junit.org) framework and can be run easily whenever the code is modified.

The user interface scripts are documented with Unix `man` pages. The source code is documented with comments that may be parsed into HTML using the `javadoc` command of the Java SDK.

3 Case study: connectivity-based partitioning of corpus callosum

The accompanying MICCAI paper [3] uses Camino for diffusion MRI data processing. The paper proposes a method to automate the partitioning of brain structures based on the anatomical connectivity of white matter fibres to anatomically distinct grey-matter regions. Grey matter is segmented manually in an atlas brain, and nonrigid registration between the atlas and each subject provides a grey matter segmentation for each subject. The connectivity is inferred from diffusion-MRI data using the probabilistic tractography algorithm that will shortly be integrated into Camino. In this section we examine the role of Camino in the production of that work.

Other open-source software was also used in the preparation of the paper. The image processing was carried out on PCs running GNU / Linux. The non-rigid registration algorithm [15] uses the The Insight Segmentation and Registration Toolkit (ITK). The open-source Insight Snap tool [16] was used to label the atlas brain and for data visualisation throughout the process.

3.1 Data processing with Camino

The process follows the Camino data pipeline, starting with the scanner data in raw binary format. The first step is to convert the data to voxel-order.

```
$ scanner2voxel -voxels 983040 -components 66
                          < control1.raw > control1.vo.Bfloat
```

We then fit the diffusion tensor using the nonlinear fitting routine. We pass the dtfit program the data, the scheme file, and the nonlinear option.

```
$ dtfit control1.vo.Bfloat c.scheme
         -nonlinear > control1.nonlinearDT.Bdouble
```

By default, `dtfit` expects single-precision float data and outputs tensor data as double-precision floats. The `modelfit` program offers more flexibility for advanced users, including specification of input and output data formats.

The probabilistic tractography algorithm allows one or two fibre orientations to be defined in each voxel. We fit a two-Gaussian model to the data.

```
$ modelfit -inputfile control1.vo.Bfloat -schemefile c.scheme
                    -inversion 22 > control1.twoDT.Bdouble
```

The inversion index 22 instructs Camino to fit the model $p = 0.5G(\mathbf{r}; \mathrm{D}_1, t) + G(\mathbf{r}; \mathrm{D}_2, t)$, where both tensors are cylindrically symmetric.

We use `voxelclassify` to tell the tractography when to use a two-fibre model.

```
$ voxelclassify -inputfile control1.vo.Bfloat
 -schemefile c.scheme -order 4 > control1.VCthresholds.Bdouble
```

We did not specify any F-test thresholds above, so the output from `voxelclassify` is the result of the F-test in each voxel. We select the thresholds with the `vcthreshselect` program.

These commands provided the input for the probabilistic tractography. The MICCAI paper describes the algorithm.

4 Future development

The Camino team are currently implementing probabilistic tractography and linear multi-fibre reconstruction routines. Other new features may be added as research in the field progresses and in response to user feedback. Pre releases of Camino are already in use at several centres around London and the UK. We expects its use to become widespread with the first full release. The latest release may be obtained from the Camino website (http://www.cs.ucl.ac.uk/research/medic/camino).

Acknowledgement

PAC is funded by the MIAS IRC, EPSRC GR/N14248/01 and by EPSRC GR/T22858/01.

References

1. Alexander, D.C.: An introduction to computational diffusion MRI: the diffusion tensor and beyond. In Weichert, J., Hagen, H., eds.: Visualization and Image Processing of Tensor Fields. Springer (2005)
2. Basser, P.J., Mattiello, J., Bihan, D.L.: MR diffusion tensor specstroscopy and imaging. Biophys J **66** (1994) 259–267
3. Cook, P., Zhang, H., Avants, B., Yushkevich, P., Alexander, D., Gee, J., Ciccarelli, O., Thompson, A.: An automated approach to connectivity-based partitioning of brain structures. In: Proc. MICCAI. (2005)
4. Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. **6** (2003) 750–757
5. Jones, D.K., Basser, P.J.: squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted MR data. MRM **52** (2004) 973–993
6. Chang, L.C., Jones, D.K., Pierpaoli, C.: Restore: Robust estimation of tensors by outlier rejection. MRM **52** (2005) 1088–1095
7. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. MRM **48** (2002) 331–340
8. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. Inverse Problems **19** (2003) 1031–1046
9. Tournier, J., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage **23** (2004) 1176–1185
10. Tuch, D.S.: Q-ball imaging. Magnetic Resonance in Medicine **52** (2004) 1358–1372
11. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. JMR **B111** (1996) 209–219
12. Parker, G.J.M., Alexander, D.C.: Probabilistic anatomical connectivity using persistent angular structure obtained from diffusion weighted imaging. Phil. Trans. Roy. Soc. Series B **360** (2005) 893–902
13. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. MRM **44** (2000) 41–50
14. Parker, G.J.M., Wheeler-Kingshott, C.A.M., Haroon, H.A.: A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. JMRI **18** (2003) 242–254
15. Avants, B.A., Schoenemann, P.T., Gee, J.C.: Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. MIA (in press)
16. Yushkevich, P.A., Piven, J., Cody, H., Ho, S., Gerig, G.: Regional structural characterization of the brain of schizophrenia patients. In: ISC/NA-MIC/MICCAI Workshop on Open-Source Software. (2005) http://hdl.handle.net/1926/20

Camino: Diffusion MRI reconstruction and processing

PA Cook, Y Bai, M G Hall, S Nedjati-Gilani, K K Seunarine and DC Alexander

Centre for Medical Image Computing, Department of Computer Science, University College London, UK. email: camino@cs.ucl.ac.uk

Abstract. Camino is an open-source, object-oriented software package for processing diffusion MRI data. Camino implements a data processing pipeline, which allows for easy scripting and flexible integration with other software. This paper summarises the features of Camino at each stage of the pipeline from the raw data to the statistics used by clinicians and researchers. The paper also discusses the role of Camino in the paper "An Automated Approach to Connectivity-based Partitioning of Brain Structures", published at MICCAI 2005.

1 Introduction

Diffusion-weighted MRI allows us to examine the characteristics of water diffusion *in vivo*. Diffusion MRI measures the particle displacement density p in each voxel of the image volume. The anisotropy of p reflects that of the tissue microstructure. Statistics of diffusion are used extensively in clinical medicine and biomedical research.

Diffusion MRI sequences typically acquire N measurements with different "diffusion weighting" gradients in each voxel. Each measurement is a sample of the Fourier transform of p a the wavenumber $\mathbf{q}$. Several reconstruction algorithms exist to fit models of p or extract useful features (such as anisotropy or fibre orientation) from the data [1]. The most widely used model is the diffusion tensor [2], which assumes a zero-mean, Gaussian p:

$$p = G(\mathbf{r}; \mathrm{D}, t) = ((4t)^3 \det(\mathrm{D}))^{-\frac{1}{2}} \exp(-\mathbf{r}^{\mathrm{T}} \mathrm{D}^{-1} \mathbf{r} (4t)^{-1}) \tag{1}$$

where D is a symmetric, second-order tensor. The log of each measurement provides a linear constraint on D:

$$\log[A^*(\mathbf{q})/A(\mathbf{0})] = \mathbf{q}^{\mathrm{T}} \mathrm{D} \mathbf{q}\,. \tag{2}$$

Since D has six independent elements, DT-MRI requires measurements with at least six independent $\mathbf{q}$ and one at $\mathbf{q} = \mathbf{0}$, though more measurements are often made to improve the robustness of the fit.

The diffusion tensor method fails when G is a poor approximation of p. Examples of this can be seen easily in the brain when two or more bundles of fibres cross within a single voxel. Several techniques exist to reconstruct multiple fibre orientations within a voxel [1]. One method is to divide the voxel into multiple compartments such that p is a mixture of zero-mean Gaussians:

$$p = \sum_{i=1}^{M} f_i G(\mathbf{r}; \mathrm{D}_i, t) \,, \tag{3}$$

where $f_1 + f_2, \ldots, +f_M = 1$.

This paper examines the image processing steps involved in the study of diffusion-MRI and shows how the open-source Camino software suite provides the capability to extract clinically useful information from raw MRI data. In Sec. 2, we introduce the data pipeline and summarise the features of Camino. In Sec 3 we discuss the role of Camino in the companion MICCAI paper "An Automated Approach to Connectivity-based Partitioning of Brain Structures" [3].

2 Camino

Camino is an object-oriented software package for processing diffusion MR images and is written entirely in Java. Users may invoke the Java programs directly or run the wrapper (bash shell) scripts provided with the software. Advanced users can construct custom scripts or extend the source code to suit their needs.

The pipeline approach to data processing is illustrated in Fig. 1. The shell scripts realise the pipeline approach and provide a stable interface for users, which means that no knowledge of Java is required to use Camino. This section examines each stage of the pipeline in detail.

2.1 Data sources

The data source can be data from a scanner or from Camino's data synthesiser. Data consists of N measurements in each voxel. Camino processes all data in voxel order, where the measurements for each voxel are stored together. This ordering facilitates the data pipeline model and allows each voxel to be processed independently. The data pipeline allows easy parallel processing and processing of large images with minimal memory consumption.

Data from scanners is not typically in voxel order, since the scanner acquires a measurement over the whole image volume, then acquires the second measurement, and so on for all N measurements. We call this "scanner order". The Camino command `scanner2voxel` rearranges data into voxel order.

```
$ scanner2voxel -components 60 -numvoxels 983040 <
   SubjectA.raw > SubjectA.vo.Bfloat
```

Camino's output is raw binary, with no header information. We therefore follow a naming convention for output, appending ".Bfloat" to the file name to identify the data format as big-endian float. The convention in Java is to read and

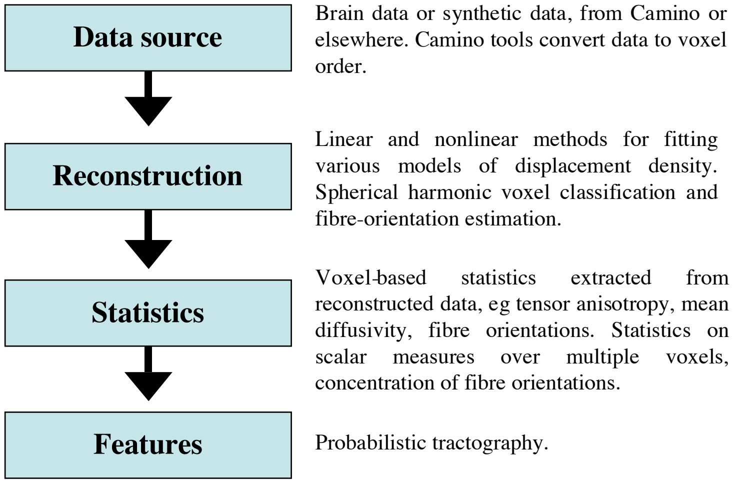


Fig. 1. The data-processing pipeline.

write big-endian data, we follow this convention in Camino. Data in little-endian format may be converted using the `shredder` program.

Every data file has an associated "scheme file", which specifies the scanner protocol used in the data acquisition. Specifically, the scheme file contains the diffusion time and wavenumber for each measurement.

Camino's data synthesizer emulates scanner sequences given a scheme file and a test function for p. Measurements can be synthesised independently for each voxel, or by bootstrap sampling from a small number of synthesised measurements. The Gaussian model (Eq. 1), and the multi-Gaussian model (Eq. 3) are implemented in Camino for use in data synthesis. Developers can add alternative models of p (eg the model proposed in [4]) by implementing the simple `ModelPDF` interface:

```
public interface ModelPDF {

    /**
     * Returns the value of the function at the specified point.
     *
     * @param x The point to sample at.
     * @return The value of the function at x.
     */
    public double at(double[] x);

    /**
     * Returns the value of the FT of the function at the
     * specified wavenumber.
```

```
     *
     * @param q The wavenumber to sample at.
     * @return The value of the FT at q.
     */
    public double ftAt(double[] q);

    /**
     * Returns a list of principal directions of the test
     * function.
     *
     * @return Array of principal directions.  The first index is the
     * principal direction number.
     */
    public double[][] getPDs();

}
```

2.2 Reconstruction

The reconstruction stage takes as input raw data and reconstructs information about p. The following command takes as input the raw data from "subject A", the associated scheme file, and fits D to the log measurements by linear regression:

```
$ dtfit SubjectA.Bfloat A.scheme > DT_A.Bdouble
```

The program outputs eight values for each voxel, in raw binary format. The first value is the error code, which is zero if the fit was successful, otherwise a nonzero code describes the nature of the error. The second value is $\log[A(\mathbf{0})]$, the estimated signal at $\mathbf{q} = \mathbf{0}$. The remaining six values are the six independent elements of D. The `dtfit` program processes a data set of $128 \times 128 \times 60$ voxels with 66 measurements per voxel in approximately two minutes on a 2.4 GHz Pentium PC.

The `dtfit` program can fit D directly to the measurements using a Levenberg-Marquardt algorithm, which offers better noise immunity [5]. Camino also supports the RESTORE method [6] for nonlinear fitting.

Camino features several high angular resolution or "multi fibre" reconstruction methods. The spherical harmonic algorithm of Alexander et al [7] detects and models non-Gaussian diffusion. The algorithm fits a nested hierarchy of spherical harmonic functions to the data and uses an F-test for deletion of variables to decide when to truncate the series. The results of the F-test identify areas of isotropic diffusion (series order 0 spherical harmonic), anisotropic Gaussian diffusion (order 2) or non-Gaussian anisotropic diffusion (order 4 or higher). The `voxelclassify` program in Camino provides voxel classification using this algorithm. Sec. 3 shows an example use of this program.

Camino implements two-Gaussian and three-Gaussian model fitting (Eq. 3). Users may optionally constrain the fitting process to improve the results. The mixing fraction f in Eq. 3 may be fixed and the fitted diffusion tensors may be constrained to cylindrical symmetry.

The persistent angular structure [8] is a special case of spherical deconvolution [9], both are implemented in the current release. These methods are computationally intensive, nonlinear inversions of the diffusion data. The Camino team is currently testing linear spherical deconvolution and the Q-ball algorithm [10] to resolve multiple fibres. Future releases will include these algorithms.

2.3 Statistics

Voxel-based statistics of the diffusion can be extracted from the reconstructed data. For tensors, two common statistics are Tr(D) and the fractional anisotropy [11]. The Camino pipeline allows users to produce these statistics from raw data by combining each stage into a single command:

```
$ scanner2voxel -components 60 -numvoxels 983040
    < SubjectA.raw | dtfit - A.scheme | fa
    -inputdatatype double > A.fa.Bdouble
```

Tr(D) is calculated in a similar manner, using the command `trd` instead of `fa`. The `dteig` program outputs the entire eigen system of the tensor. The `shredder` program extracts volumes from composite data sets. All of these programs handle one, two or three tensors per voxel. Statistics on populations of voxels include the mean and variance of scalar measures such as anisotropy, or the concentration of fibre orientations.

Camino also provides voxel and population statistics from non-tensor data, such as the output of PAS-MRI or spherical deconvolution. The `sfpeaks` program extracts peaks of spherical functions and the Hessian at those peaks for use in tractography [12].

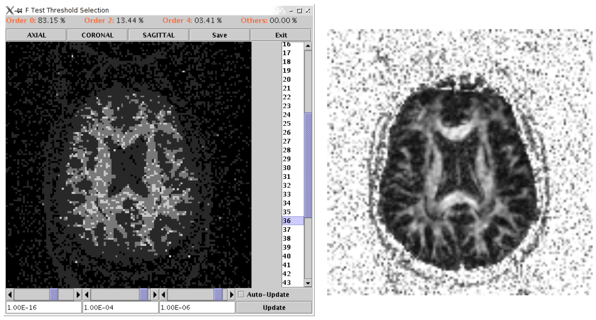
2.4 Image features

Deterministic and probabilistic tractography [13, 14] are currently being integrated into the Camino suite. These algorithms will produce streamlines or connectivity maps for connectivity analysis.

2.5 Visualisation

Camino's primary purpose is to provide a pipeline-based suite of data processing algorithms. All of Camino's output data structure is documented kept as simple as possible to facilitate integration with visualisation software. Visualisation features may be added to future releases of Camino as another stage in the pipeline, but are not a priority.

Camino currently contains only one visualisation tool, which is used to apply the F-test for spherical harmonic voxel classification. These thresholds can be determined interactively (Fig. 2) by piping the output of `voxelclassify` into the `vcthreshselect` program.


F Test Threshold Selection

Order 0: 83.15 % Order 2: 13.44 % Order 4: 03.41 % Others: 00.00 %

1.00E-16 1.00E-04 1.00E-06

Fig. 2. The `vcthreshselect` program (left) and a slice from an anisotropy map (right).

2.6 Testing and documentation

The modular nature of Camino facilitates unit testing. Camino tests are written using the open source JUnit (http://www.junit.org) framework and can be run easily whenever the code is modified.

The user interface scripts are documented with Unix `man` pages. The source code is documented with comments that may be parsed into HTML using the `javadoc` command of the Java SDK.

3 Case study: connectivity-based partitioning of corpus callosum

The accompanying MICCAI paper [3] uses Camino for diffusion MRI data processing. The paper proposes a method to automate the partitioning of brain structures based on the anatomical connectivity of white matter fibres to anatomically distinct grey-matter regions. Grey matter is segmented manually in an atlas brain, and nonrigid registration between the atlas and each subject provides a grey matter segmentation for each subject. The connectivity is inferred from diffusion-MRI data using the probabilistic tractography algorithm that will shortly be integrated into Camino. In this section we examine the role of Camino in the production of that work.

Other open-source software was also used in the preparation of the paper. The image processing was carried out on PCs running GNU / Linux. The non-rigid registration algorithm [15] uses the The Insight Segmentation and Registration Toolkit (ITK). The open-source Insight Snap tool [16] was used to label the atlas brain and for data visualisation throughout the process.

3.1 Data processing with Camino

The process follows the Camino data pipeline, starting with the scanner data in raw binary format. The first step is to convert the data to voxel-order.

```
$ scanner2voxel -voxels 983040 -components 66
                          < control1.raw > control1.vo.Bfloat
```

We then fit the diffusion tensor using the nonlinear fitting routine. We pass the dtfit program the data, the scheme file, and the nonlinear option.

```
$ dtfit control1.vo.Bfloat c.scheme
         -nonlinear > control1.nonlinearDT.Bdouble
```

By default, `dtfit` expects single-precision float data and outputs tensor data as double-precision floats. The `modelfit` program offers more flexibility for advanced users, including specification of input and output data formats.

The probabilistic tractography algorithm allows one or two fibre orientations to be defined in each voxel. We fit a two-Gaussian model to the data.

```
$ modelfit -inputfile control1.vo.Bfloat -schemefile c.scheme
                    -inversion 22 > control1.twoDT.Bdouble
```

The inversion index 22 instructs Camino to fit the model $p = 0.5G(\mathbf{r}; \mathrm{D}_1, t) + G(\mathbf{r}; \mathrm{D}_2, t)$, where both tensors are cylindrically symmetric.

We use `voxelclassify` to tell the tractography when to use a two-fibre model.

```
$ voxelclassify -inputfile control1.vo.Bfloat
 -schemefile c.scheme -order 4 > control1.VCthresholds.Bdouble
```

We did not specify any F-test thresholds above, so the output from `voxelclassify` is the result of the F-test in each voxel. We select the thresholds with the `vcthreshselect` program.

These commands provided the input for the probabilistic tractography. The MICCAI paper describes the algorithm.

4 Future development

The Camino team are currently implementing probabilistic tractography and linear multi-fibre reconstruction routines. Other new features may be added as research in the field progresses and in response to user feedback. Pre releases of Camino are already in use at several centres around London and the UK. We expects its use to become widespread with the first full release. The latest release may be obtained from the Camino website
(http://www.cs.ucl.ac.uk/research/medic/camino).

Acknowledgement

PAC is funded by the MIAS IRC, EPSRC GR/N14248/01 and by EPSRC GR/T22858/01.

References

1. Alexander, D.C.: An introduction to computational diffusion MRI: the diffusion tensor and beyond. In Weichert, J., Hagen, H., eds.: Visualization and Image Processing of Tensor Fields. Springer (2005)
2. Basser, P.J., Mattiello, J., Bihan, D.L.: MR diffusion tensor specstroscopy and imaging. Biophys J **66** (1994) 259–267
3. Cook, P., Zhang, H., Avants, B., Yushkevich, P., Alexander, D., Gee, J., Ciccarelli, O., Thompson, A.: An automated approach to connectivity-based partitioning of brain structures. In: Proc. MICCAI. (2005)
4. Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. **6** (2003) 750–757
5. Jones, D.K., Basser, P.J.: squashing peanuts and smashing pumpkins: How noise distorts diffusion-weighted MR data. MRM **52** (2004) 973–993
6. Chang, L.C., Jones, D.K., Pierpaoli, C.: Restore: Robust estimation of tensors by outlier rejection. MRM **52** (2005) 1088–1095
7. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. MRM **48** (2002) 331–340
8. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. Inverse Problems **19** (2003) 1031–1046
9. Tournier, J., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage **23** (2004) 1176–1185
10. Tuch, D.S.: Q-ball imaging. Magnetic Resonance in Medicine **52** (2004) 1358–1372
11. Basser, P.J., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. JMR **B111** (1996) 209–219
12. Parker, G.J.M., Alexander, D.C.: Probabilistic anatomical connectivity using persistent angular structure obtained from diffusion weighted imaging. Phil. Trans. Roy. Soc. Series B **360** (2005) 893–902
13. Basser, P.J., Pajevic, S.: Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. MRM **44** (2000) 41–50
14. Parker, G.J.M., Wheeler-Kingshott, C.A.M., Haroon, H.A.: A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. JMRI **18** (2003) 242–254
15. Avants, B.A., Schoenemann, P.T., Gee, J.C.: Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. MIA (in press)
16. Yushkevich, P.A., Piven, J., Cody, H., Ho, S., Gerig, G.: Regional structural characterization of the brain of schizophrenia patients. In: ISC/NA-MIC/MICCAI Workshop on Open-Source Software. (2005) http://hdl.handle.net/1926/20