DICOM and NIfTI Data Standards
in R

Brandon Whitcher PhD CStat

Mango Solutions
London, United Kingdom
www.mango-solutions.com
bwhitcher@mango-solutions.com
@MangoImaging

12 June 2012 – useR!2012 Tutorial

MANGOSOLUTIONS
data analysis that delivers

Outline

1. **Summary**

2. **Digital Imaging and Communications in Medicine**

3. **Neuroimaging Informatics Technology Initiative**

4. **Data Visualization**

5. **DICOM-to-NIfTI Conversion**

6. **Conclusions**

- Content presented here provides users with the skills to manipulate DICOM / ANALYZE / NIfTI files in **R**.
  - Real-world data sets are used to illustrate the basic functionality of **oro.dicom** and **oro.nifti**.
  - S4 classes "nifti" and "anlz" enable further statistical analysis in **R** without losing contextual information from the original ANALYZE or NIfTI files.
  - Images in the metadata-rich DICOM format may be converted to NIfTI semi-automatically using as much information from the DICOM files as possible.
  - Basic visualization functions, are provided for "nifti" and "anlz" objects.
  - The **oro.nifti** package allows one to track every operation on a nifti object in an **XML**-based audit trail.

Alternatives to **oro.dicom** and **oro.nifti**

- There are several **R** packages that are able to access the DICOM / ANALYZE / NIfTI data formats:
  - **AnalyzeFMRI** (Bordier et al. 2009)
  - **fmri** (Polzehl and Tabelow 2007)
  - **Rniftilib** (Granert 2010)
  - **tractor.base** (Clayden 2010)

Question #1

What are the (dis)advantages to having a single **R** package that performs input / output for medical imaging data?

Question #2

Should **R** packages be discouraged from writing output in formats other than ANALYZE or NIfTI?

The DICOM "Standard"

- The DICOM "standard" for data acquired using a clinical imaging device is very broad and complex.
  - Each DICOM-compliant file is a collection of fields organized into two two-byte sequences (group,element) that are represented as hexadecimal numbers and form a tag.
  - The (group,element) combination establishes what type of information is forthcoming in the file.
- There is no fixed number of bytes for a DICOM header.
- The final (group,element) tag should be the *PixelData* tag (7FE0,0010), such that all subsequent information is related to the image.

DICOM

Digital Imaging and Communications in Medicine
`http://medical.nema.org`

The Structure of a DICOM File

```
Data element with explicit VR of OB, OF, OW, SQ, UT or UN:

+-----------------------------------------------------------+
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
+----+----+----+----+----+----+----+----+----+----+----+----+
|<Group-->|<Element>|<VR----->|<0x0000->|<Length----------->|<Value->

Data element with explicit VR other than as shown above:

+---------------------------------------+
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
+----+----+----+----+----+----+----+----+
|<Group-->|<Element>|<VR----->|<Length->|<Value->

Data element with implicit VR:

+---------------------------------------+
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
+----+----+----+----+----+----+----+----+
|<Group-->|<Element>|<Length----------->|<Value->
```

- Byte ordering for a single (group,element) tag in the DICOM standard.
- Explicit VRs store the VR as text characters in two bytes.

Reading DICOM Files

readDICOMFile()

Accessing information stored in a single DICOM file

- The resulting object is a list with two elements: the DICOM header (hdr) and the DICOM image (img).
- Header information is organized in a data frame with six columns and an unknown number of rows.
- First five columns taken from DICOM header information (group, element, code, length and value) or inferred (name).
- (group,element) values are stored as character strings – not hexadecimal numbers.

readDICOM()

Accessing multiple DICOM files in a single directory or directory tree

create3D() and create4D()

Create arrays from DICOM headers / images

- Minimum input = "dcm" structure
- *PixelData* may be read on-the-fly
- Siemens MOSAIC format allowed
- 4D volumes may require additional information
  - nslices = ?
  - ntimes = ?

The ANALYZE Format

- The ANALYZE format was originally developed in conjunction with an image processing system (of the same name) at the Mayo Foundation.
- A common version of the format, although not the most recent, is called ANALYZE 7.5.
- An ANALYZE 7.5 format image is comprised of two files, the ".hdr" and ".img" files, that contain information about the acquisition and the acquisition itself, respectively.

The ANALYZE Format

.hdr = 348-byte binary file of header information
.img = binary flat file of images (multi-dimensional array)

The NIfTI-1 Format

- The NIfTI-1 data format is almost identical to the ANALYZE format, but offers a few improvements:
  - Merging of the header and image information into one file ".nii"
  - Re-organization of the 348-byte fixed header into more relevant categories
  - Possibility of extending the header information
- Discussions have begun on the NIfTI-2 data format.

NIfTI

Neuroimaging Informatics Technology Initiative
`http://nifti.nimh.nih.gov`

- "nifti"
  - Inherits from class "array"
  - Slots contain NIfTI header information
  - Basic methods: show(), nifti(), is.nifti(), as(<obj>, "nifti")
  - Input / output: readNIfTI(), writeNIfTI()
  - Slot access: pixdim(), qform(), sform(), descrip(), aux.file(), audit.trail()
  - Additional classes: "niftiExtension" and "niftiExtensionSection"
- "anlz"
  - Inherits from class "array"
  - Slots contain ANALYZE header information

Visualization

- **oro.nifti** offers three functions for visualization
  - image() = overloaded function for "anlz", "array" and "nifti" classes
  - overlay() = extension of image() with x and y input parameters
  - orthographic() = mid-axial, mid-sagittal, mid-coronal views

Interactive Visualization

FSLView

`http://www.fmrib.ox.ac.uk/fsl/fslview/`

MRIcron

`http://www.cabiatl.com/mricro/mricron/`

VolView

`http://www.kitware.com/products/volview.html`

DICOM-to-NIfTI Conversion

- **oro.dicom** and **oro.nifti** were designed to use as much information as possible from the metadata-rich DICOM format and apply that information in the construction of the NIfTI data volume.
  - Read in a single series using dicomSeperate()
  - dicom2nifti() converts the list of DICOM images into a multidimensional "nifti" object
  - dicom2analyze() converts the list of DICOM images into a multidimensional "anlz" object
  - Additional scripting in **R** is required to deal with multiple series

- Content presented here provides users with the skills to manipulate DICOM / ANALYZE / NIfTI files in **R**.
  - Real-world data sets are used to illustrate the basic functionality of **oro.dicom** and **oro.nifti**.
  - S4 classes "nifti" and "anlz" enable further statistical analysis in **R** without losing contextual information from the original ANALYZE or NIfTI files.
  - Images in the metadata-rich DICOM format may be converted to NIfTI semi-automatically using as much information from the DICOM files as possible.
  - Basic visualization functions, are provided for "nifti" and "anlz" objects.
  - The **oro.nifti** package allows one to track every operation on a nifti object in an **XML**-based audit trail.

- `http://rigorousanalytics.blogspot.com`
- `https://r-forge.r-project.org/projects/ rigorous`
- `rigorous-dicom@lists.r-forge.r-project.org`
- `rigorous-nifti@lists.r-forge.r-project.org`
- `rigorous-dcemris4@lists.r-forge.r-project.org`

Volume 44 in the *Journal of Statistical Software*

Special volume on "Magnetic Resonance Imaging in R"

- 13 articles on fMRI, DTI, DCE-MRI, etc.
- `www.jstatsoft.org/v44`

DICOM and NIfTI Data Standards in R

Summary

DICOM

NIfTI

Visualization

Conversion

End

Bibliography

Bibliography I

C. Bordier, M. Dojat, and P. Lafaye de Micheaux. **AnalyzeFMRI**: an **R** package to perform statistical analysis on fmri datasets. 2009. URL `http://www.biostatisticien.eu/AnalyzeFMRI/`. Software: **R** Package, **AnalyzeFMRI**, version 1.1-12.

J. Clayden. **_tractor.base_**_: A Package for Reading, Manipulating and Visualising Magnetic Resonance Images_, 2010. URL `http://CRAN.R-project.org/package=tractor.base`. **R** package version 1.5.0.

O. Granert. **_Rniftilib_**_: R Interface to_ **NIFTICLIB** _(V1.1.0)_, 2010. URL `http://CRAN.R-project.org/package=Rniftilib`. **R** package version 0.0-29.

J. Polzehl and K. Tabelow. **fmri**: A package for analyzing fmri data. *RNews*, 7(2):13–17, 2007. URL `http://www.r-project.org/doc/Rnews/Rnews_2007-2.pdf`.

R Package: **oro.dicom**

- **Title:** Rigorous - DICOM Input / Output
- **Description:** Data input/output functions for data that conform to the Digital Imaging and Communications in Medicine (DICOM) standard, part of the Rigorous Analytics bundle.
- **Depends:** R (>= 2.13.0), utils
- **Suggests:** hwriter, oro.nifti (>= 0.2.9)
- **License:** BSD
- **URL:** http://rigorousanalytics.blogspot.com

R Package: **oro.nifti**

- **Title:** Rigorous - NIfTI+ANALYZE+AFNI Input / Output
- **Description:** Functions for the input/output and visualization of medical imaging data that follow either the ANALYZE, NIfTI or AFNI formats. This package is part of the Rigorous Analytics bundle.
- **Depends:** R (>= 2.13.0), bitops, graphics, grDevices, methods, utils
- **Suggests:** XML
- **Imports:** splines
- **Enhances:** fmri
- **License:** BSD
- **URL:** http://rigorousanalytics.blogspot.com