

DOI:10.3969/j.issn.1005-202X.2016.12.019

医学影像物理

磁共振成像数据自动整理工具箱的研制

裴浩男, 曹卫芳, 郭永新, 崔栋, 焦青
泰山医学院放射学院, 山东 泰安 271016

【摘要】目的:在对批量DICOM格式的磁共振成像数据进行处理之前,需要对数据文件夹进行重命名、统一编号以及格式转换等整理工作,手工进行该项工作耗时、耗力且易出错,为了解决此困难,本研究研制出一个磁共振成像数据自动整理工具箱。**方法:**在Windows环境下,采用MATLAB的图形用户界面进行开发,并采用简单函数对数据进行读写操作,调用内置程序进行数据处理。**结果:**该工具箱可自动完成对磁共振成像的DICOM数据文件名的重定、统一编号、格式转换与分类放置的功能。**结论:**该工具箱界面友好、易于操作,可极大地提高科研工作效率。

【关键词】DICOM数据;MATLAB;自动整理工具箱;磁共振成像

【中图分类号】TP315

【文献标志码】A

【文章编号】1005-202X(2016)12-1267-05

Development of auto arrangement toolbox of magnetic resonance imaging data

PEI Haonan, CAO Weifang, GUO Yongxin, CUI Dong, JIAO Qing
Department of Radiology, Taishan Medical University, Tai'an 271016, China

Abstract: Objective Before processing huge magnetic resonance imaging (MRI) data stored in digital imaging and communication in medical (DICOM) format, some arrangement steps such as renaming and sorting data files, and data format conversation usually are performed manually, which is time-consuming and error-prone. An auto arrangement toolbox of the MRI data was developed to solve the above difficulties. **Methods** In Windows environment, the graphical user interface of MATLAB was used for the development. Simple functions and built-in procedures were used for reading, writing and processing of the data. **Results** Renaming and sorting DICOM data files, data format conversation and classification of data files were completed automatically by using the proposed toolbox. **Conclusion** With a friendly interface, the toolbox is easy to be operated, greatly improving the work efficiency.

Keywords: digital imaging and communication in medical data; MATLAB; auto arrangement toolbox; magnetic resonance imaging

前言

目前,磁共振成像(Magnetic Resonance Imaging, MRI)的重建数据都采用DICOM标准(Digital Imaging and Communication in Medical)进行存储和传输^[1]。DICOM标准是专用于图像存储和传输的标准协议,用来定义图像及其相关信息的组成格式和交换方法,进而完成图像的输入与输出,促进设备的兼容性。与

其他图像文件不同的是,DICOM格式的图像文件不仅包含图像数据,还包含许多与图像有关的文本信息,如病人基本资料、检验基本资料以及描述图像的基本参数等^[2]。这些文本信息是医学图像数据库中重要的数据组成说明部分,但由于数据被封装在文件中,一般需要特定软件进行读取^[1]。在实际应用中,DICOM数据一般存放在以被试名字命名的文件夹中。

在对MRI原始DICOM数据进行后处理时,存在一些问题:(1)有的DICOM文件名被设成了被试的中文名字(汉字),这给后续数据处理带来困难,因为目前多数MRI数据处理软件无法识别汉字,这会导致无法进行后续处理,因此必须将文件夹的汉字名称转换为英文字母;(2)当处理大量数据时,需要统一编号以防止数据遗漏丢失;(3)通常研究者需要把

【收稿日期】2016-10-30

【基金项目】国家自然科学基金(81371531);山东省医药卫生科技发展计划项目(2016WS0603);山东省高等学校科技计划项目(J14LK55);泰安市科技发展计划(2015NS2073)

【作者简介】裴浩男,研究方向:生物医学信号处理,E-mail: 1073741355@qq.com

【通信作者】焦青,博士,教授,研究方向:生物医学信号处理,磁共振成像技术,E-mail: bingbao17@163.com

同一次采集的 MRI 不同序列的数据分别处理,因此需要先把 DICOM 文件进行格式转换,如 Nifti 格式,并将不同序列的文件分开放置^[3]。Nifti 格式的文件包含了头文件及 DICOM 文件中各序列的图像信息,如氢元素的纵向弛豫时间(T_1)、横向弛豫时间(T_2)、血氧水平依赖(Blood Oxygen Level Dependent, BOLD)和弥散张量成像(Diffusion Tensor Imaging, DTI)等, Nifti 格式的文件使用以往规范中的一些未使用或很少使用的字段以存储新的信息,例如将脑图像的左右方向存储其中^[4]。Nifti 格式是当前常用的神经影像学软件研究的默认格式,如 FMRIB 软件库(FMRIB Software Library, FSL)^[5]、统计参数图(Statistical Parametric Mapping, SPM)^[6]和功能神经影像分析软件(Analysis of Functional NeuroImages, AFNI)^[7],我国研究者开发的静息态 fMRI 数据处理工具箱(Resting-State fMRI Data Analysis Toolkit)^[8],静息态 fMRI 数据处理助手(Data Processing Assistant for Resting-State fMRI)^[9],以及脑张量图像分析工具包 PANDA(Pipeline for Analyzing brain Diffusion images)^[10]。另外,许多图像查看与分析软件(如 3D Slicer、ImageJ 和 OsiriX)以及其他新兴软件(如 Nibabel)同样支持使用 Nifti 格式^[4]。有时,使用者并不需要处理全部序列的文件,如有的工作只需要处理 DTI 数据^[11]。故需要将 Nifti 格式文件进行分类。DICOM 数据文件名的重定、统一编号、格式转换与分类放置这些工作通常采用手工进行,耗时、耗力且容易出错。

为了解决上述问题,本研究在 MATLAB 环境下开发了一个 MRI 数据整理工具箱(NMRI Data Arrangement Toolbox, NDAT),可自动完成对磁共振成像仪采集的 DICOM 数据文件名的重定、统一编号、格式转换与分类放置的功能,提高数据分析人员的工作效率。

1 工具箱框架

NDAT 工具箱可在 Windows 系统(Win 7.0)与 Linux 系统(Centos 7.0)下运行,运行内存 2 GB 以上即可,存储空间根据所处理数据量不同而不同,单纯存放代码文件只需 25 MB。工具箱在图形用户界面设计环境(Graphics User Interface Design Environment, GUIDE)中开发完成,GUIDE 是一个专用于 GUI 程序设计的快速开发环境,开发者可对各种 GUI 控件的属性进行操作,从而设计出符合要求的图形用户界面,已经获得广泛应用^[12]。工具箱主界面如图 1 所示,包括菜单栏、文件路径栏、选项栏及功能栏。菜单栏

内含 3 个小功能:(1)语言(Language)功能,可供操作者选择的界面语言(汉语或英语);(2)复位(Reset)功能,可清空主界面中的文本框信息;(3)帮助(Help),提供工具箱的使用说明。文件路径栏可供操作者选择文件的输入路径与输出路径并将其显示在文本框中。选项栏中,可输入项目名的缩写(Project Abb.),这是考虑到用户可能会处理多个课题的数据,为了方便记忆与区分,由操作者确定欲处理数据的项目主题词,如癫痫或抑郁症;小组名(Group Name)指的是该项目下的被试小组名,如健康人组或病人组;初始序号(Initial Number)是操作者自主选择排序的起始序号。功能栏即主要功能选择区,其中排序(Listing)是指将读入的文件夹排序,并在文件夹名前加上相对应的序号;重命名(Change Name)功能将每个文件夹名中的中文以相应的英文替换形成新名字;Dcm2nii 功能将 DICOM 文件转换为 Nifti 格式文件;分类(Classification)功能将生成的 Nifti 格式文件按照不同 MRI 采集序列分类放置。

工具箱操作简单,操作者只需先选择文件加载工具箱所在路径,在 MATLAB 命令栏输入 NDAT 回车就可出现如图 1 所示的界面,选择输入与输出路径,输入项目主题词(Project Abb.)、小组名(Group Name)和初始序号(Initial Number),然后在功能栏中选择需要执行的具体功能,最后点击运行(Run)即可执行。

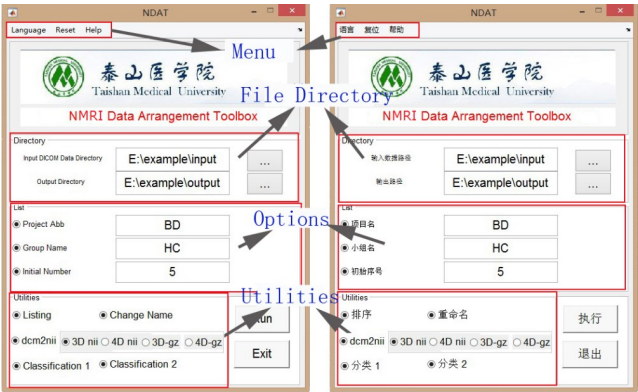


图 1 工具箱主界面
Fig.1 Main Interface of Toolbox

2 功能实现

2.1 文件夹重命名的实现

进行了相关功能选项之后,文件夹的改名过程如图 2 所示,绿色滚动条显示改名进程。图 3 为改名后的结果,其中 a 为输入路径;b 为待处理文件夹,每一个文件夹中有一组被试者数据,内含 DICOM 文件,文件夹名即为被试者姓名;c 为输出文件路径,其

中包含项目名(如c中“BD”)与小组名(如c中“HC”)两个分级文件夹;d为处理后文件,每个文件夹中有一组被试者数据。

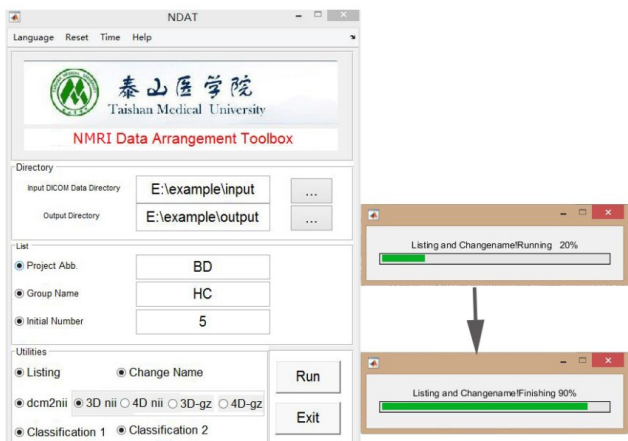


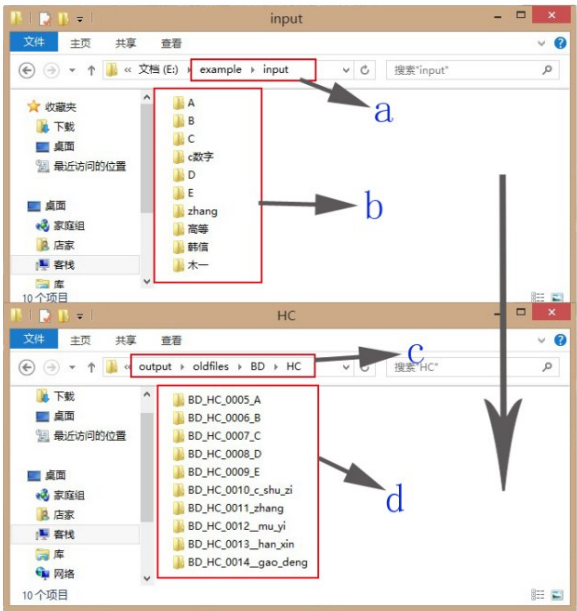
图2 文件的重命名
Fig.2 Renaming Files

绝大多数DICOM数据的文件夹以被试者姓名命名(包含中文或英文,或二者皆有),少数还会有使用者附加的序号等标识符,比如‘c’、‘0003’。工具箱的“Change Name”功能可实现将原有文件名中的汉字转换为对应汉语拼音,只对汉字进行处理,其他字符不变。比如源文件夹名为“c数字”,转换后新文件名为“c_shu_zi”。重命名时如果选中“Project Abb.”与“Group Name”并输入字段,则输出文件名中会以项目名与小组名作为前缀,比如“BD_HC_0010_c_shu_zi”(图3的b和d),这两个前缀也是将DICOM文件分类的主要依据,清晰明确,使用者只需看前缀即可明白这一数据属于哪一组被试者。

工具箱中拥有一个汉字与相应拼音对应的function文件,在处理时将文件名以字符串(记为a)的形式读入,通过function文件完成汉字转换为拼音的工作,并将新名称以字符串(记为b)形式返回,然后以字符串b为名创建新文件夹,随即将文件夹a中的文件全部移动到新文件夹b中并删除文件夹a,达到改名的目的。

2.2 自动排序的实现

自动排序是将文件进行自动编号,在原文件名之前加入数字序号,包含以下3部分:(1)“Listing”的基础功能是为文件夹排序,以一递增,直到最后一个。排序时,先用dir函数读取源文件并记录文件数量,然后进入for循环开始重复执行重命名(只是在原文件名前添加序号组成新字符串作为文件名),但每次排序序号加1。一次循环只对一个文件夹进行修



a: Input path; b: Pending file folders; c: Output path; d: Worked files

图3 文件的排序
Fig.3 Sorting Files

改;(2)若操作者想要从特定序号(比如003)开始,只需选中“Initial Number”并添加特定序号即可。此功能是为同一批数据存放于不同文件时提供(比如同一批采集的癫痫MRI数据,但因为某种原因被存放在不同的文件夹中),可继承前一组数据尾号加1(即用户所提供序号)继续排序,不影响文件数目的判定,同时方便操作者确定是否为同一批数据,并准确定位到这一系列数据中的某一份数据。当选中该功能时,get函数会读取文本框中信息作为基础数值,然后进入循环,此时会将每次的数据编号加上这一基础数值作为真正序号,同时执行字符串的修订生成新名替代原名;(3)若多次读入新数据放入同一文件夹中,此时将以原有文件夹数量与使用者提供序号(若有的话)之和为初始序号。在读入新数据之前,会对目标文件夹进行扫描,用dir函数读取信息确定是否为空,若不为空则记录内部文件数量,将其作为基础数值并加入每一次的排序循环中,达到同一文件夹分批存储数据,却拥有连续序号的目的。若想使用此功能,只需选中并在“Initial Number”中提供初始编号即可。

2.3 格式转换的实现

格式转换是将每个被试的DICOM文件转换为相应的Nifti格式文件,包括压缩Nifti格式文件(后缀为“.gz”)和非压缩Nifti格式文件。若选中“3D”或“4D”选项,则分别表示输出三维或四维(三维加时间)的Nifti格式文件,保存为四维数据格式可减少输出的Nifti格式文件数量(图4)。

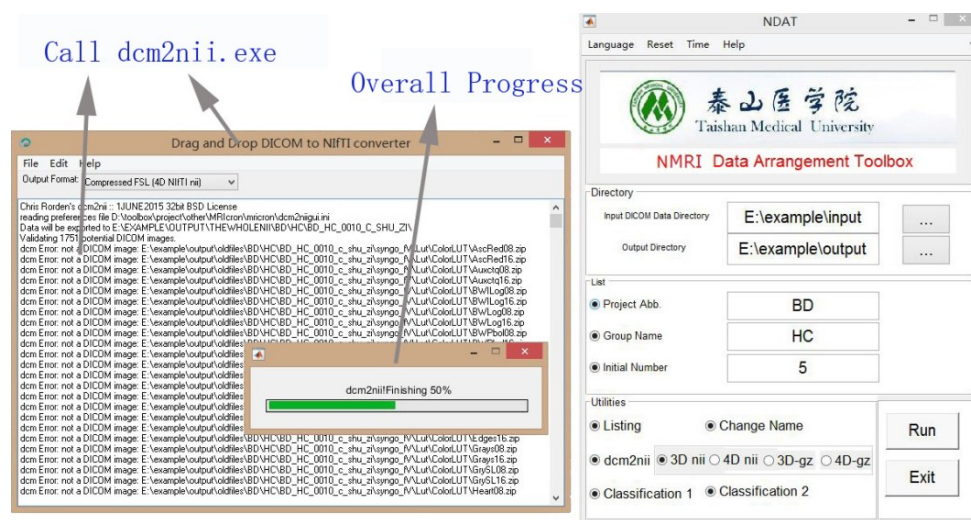


图4 格式转换

Fig.4 Format Conversion

运行时,由dir函数读取并记录本次所要处理的所有文件夹的名字等信息,以for循环进行连续转换。先读取输出路径文件夹中的文件名字,将本次循环时所要处理的文件夹的名字与其作比较,若有相同的名字则说明已经存在同名文件若存在,则跳过,以保护原有数据同时避免重复执行;若不存在,则使用system函数调用“dcm2niigui.exe”程序进行格式转换,并输出到目标路径下。在转换结束后开启下一次处理。

2.4 Nifti 格式文件分类的实现

通过读取DICOM文件的信息,依据不同的成像序列,将生成的Nifti格式文件以不同的类别分开放置于同一组下不同文件夹中。本工具箱提供两种分类方式:(1)分类1(Classification1):先将Nifti格式文件以“被试者”为类别分开存放,然后依据不同的成像序列将其分类存放。经此分类后,每个被试者文件夹下包含此人所有序列的Nifti格式文件。在运行时,首先程序会读取每个被试者的Nifti格式文件,然后读取输出文件夹内的所有文件名,并依次与本次循环时所要处理的被试者文件夹的名字作比较,若两个名字相同则说明此被试者文件夹已存在(若存在则跳过,以保护原有数据同时避免重复执行);若两个名字不同则此被试者文件夹不存在,此时立刻寻找相对应的DICOM文件,并用dicominfo函数读取DICOM文件信息,利用其中的ProtoCoName属性作为分类的依据,用copyfile函数复制文件并分类存放;(2)分类2(Classification2):先将Nifti格式文件以相应不同的成像序列分类放置,再将Nifti格式文件以被试者为类别放开放置。这种分类方式可以把不同序列分开,每个序列文件夹下包含所有被试者的该序列的文件。操作者可以根据需要读取某一序列所有人的数据,或其中某一人数据。代码

实现方式与分类1基本相同,区别在于前者先以被试者分类,后者先以序列分类。

2.5 帮助文档

在目录栏中列举了程序所有涉及的功能,通过点击选择目录,会在右侧的说明栏中给出说明。选中“分类列表(Category List)”选项,右侧自动出现解释(图5)。

3 讨论

本研究为了解决MRI数据处理工作中的实际困难,采用MATLAB语言下的GUI开发了一个自动处理工具箱,该工具箱可以实现MRI数据DICOM文件的重命名、排序、格式转换与分类功能。MATLAB工具箱实际上是具有一定功能,可解决各类问题的一系列MATLAB函数^[13]。将工具箱加载入MATLAB路径下,即可通过输入工具箱名字调用使用,准确实现特定功能而不用频繁编写代码或打开相应M文件。

GUI由窗口、菜单、对话框等各种元素组成^[14]。用户可以借助鼠标键盘操作,指挥后台程序实现功能。GUI是有别于VC界面和Java界面的一种新型界面开发方式。对于不想编写大量VC代码的科研人员来件,GUI无疑是一个更好的选择,GUI既能嵌入已有程序,又能把结果以人机交会的方式呈现给操作者。操作者不需要知道代码的具体内容,只要了解操作步骤即可很方便地操作界面^[15]。

该工具箱具有以下优点:(1)可完成批量文件夹的自动改名与编号,避免手动改名造成的失误;(2)将DICOM文件进行自动转换成Nifti格式文件,并自动按照序列和按照被试进行分类保存,方便数据处理人员按照不同的目的对数据进行后处理,比如有时使用者仅想提取BOLD-fMRI数据,则此时可以提



图5 帮助文档
Fig.5 Help Document

取BOLD文件夹里面的所有数据进行后续处理;(3)系统上述功能均可以在进行相关选项后,一键完成,实现完全自动处理,明显减少MRI数据处理人员的负担;(4)程序所涉及代码不会对电脑系统产生影响,最大程度保证程序可靠性与安全性。程序默认数据批量处理,可重复读取新数据,放入同一目标路径文件夹,不会对文件夹下原有数据进行任何修改。同时所有操作皆在备份数据中进行,不会对使用者提供的原始数据造成影响,若遇到断电等外部故障时不会导致原始数据损毁或丢失。

其他软件,如Mricron软件,同样可以完成DICOM文件的格式转换,但是Mricron软件会将生成的新数据,把不同序列(如T₁,DTI等)的Nifti格式文件都存放在原来的DICOM文件夹中,造成原始数据与新产生的数据混杂在一起,未加区分,需要进行手工分类,而纷杂错乱的数据也可能会被错误提取使用,这无疑会给操作者添加额外的工作负担。而本研究开发的工具箱则可以在Mricron软件的基础上提供格式转换,同时将转换后的数据按照序列和按照组别进行分类存放,极大地方便了后续的数据处理。

除MRI数据外,本工具箱还可以对其他DICOM格式的神经影像数据文件进行整理,另外,该工具箱不仅可以提高科研工作的效率,还可以作为教学工具,对初学者进行数据处理培训,因此本工具箱具有较高的应用价值与良好的应用前景。

【参考文献】

[1] 徐潘辉,林峰. DICOM医学数字图像格式与BMP通用图像格式转换软件的设计与实现[J]. 医学设备信息, 2006, 21(3): 1-5.
XU P H, LIN F. Design and realization of the software in transforming DICOM digitalmedical image format into BMP general image format [J]. Information of Medical Equipment, 2006, 21(3): 1-5.
[2] 尤超. DICOM文件头信息读取及解析[J]. 微型电脑应用, 2012, 28(9): 39-41.

YOU C. DICOM file header information read and analysis [J]. Microcomputer Application, 2012, 28(9): 39-41.
[3] CAO W F, LUO C, ZHU B, et al. Resting-state functional connectivity in anterior cingulate cortex in normal aging [J]. Front Aging Neurosci, 2014, 6: 280.
[4] LAROBINA M, MURINO L. Medical image file format FAQ[J]. J Digit Imaging, 2014, 27(2): 200-206.
[5] JENKINSON M, BECKMANN C F, BEHRENS T E, et al. FSL [J]. Neuroimage, 2012, 62(2): 782-790.
[6] FRISTON K J, HOLMES A P, WORSLEY K J, et al. Statistical parametric maps in functional imaging: a general linear approach [J]. Hum Brain Mapp, 1995, 2: 189-210.
[7] COX R W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages [J]. Comput Biomed Res, 1996, 29(3): 162-173.
[8] SONG X W, DONG Z Y, LONG X Y, et al. REST: a toolkit for resting-state functional magnetic resonance imaging data processing[J]. PLoS One, 2011, 6(9): e25031.
[9] YAN C G, ZANG Y F. DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI[J]. Front Syst Neurosci, 2010, 4: 13.
[10] CUI Z, ZHONG S, XU P, et al. PANDA: a pipeline toolbox for analyzing brain diffusion images [J]. Front Hum Neurosci, 2012, 7: 42.
[11] WON Y I, CHUNG C K, KIM C H, et al. White matter change revealed by diffusion tensor imaging in gliomas[J]. Brain Tumor Res Treat, 2016, 4(2): 100-106.
[12] 王巧花,叶平,黄民. 基于MATLAB的图形用户界面(GUI)设计[J]. 煤矿机械, 2005, 26(3): 60-62.
WANG Q H, YE P, HUANG M. Design of GUI base on MATLAB [J]. Coal Mine Machinery, 2005, 26(3): 60-62.
[13] PIANOSI F, SARRAZIN F, WAGENER T. A Matlab toolbox for global sensitivity analysis[J]. Environmental Modelling & Software, 2015, 70(C): 80-85.
[14] 刘卫国. MATLAB程序设计与应用[M]. 第2版. 北京: 高等教育出版社, 2006: 25-70.
LIU W G. MATLAB program design and application[M]. 2nd ed. Beijing: Higher Education Press, 2006: 25-70.
[15] 王玉林,葛蕾,李艳斌. 新型界面开发工具: MATLAB/GUI[J]. 无线电通信技术, 2008, 34(6): 50-52,
WANG Y L, GE L, LI Y B. New-style interface development tool MATLAB/GUI [J]. Radio Communications Technology, 2008, 34 (6): 50-52.

(编辑:谭斯允)