

文章编号 :1000-582X(2007)09-0126-04

DICOM 医学图像格式转换的 VC++ 实现

彭承琳 陈 诚 陈园园

(重庆大学 生物工程学院 重庆 400030)

摘 要 :DICOM3.0 是医学影像存储和传输的国际标准 ,它是随着图像化、计算机化的医疗设备的普及和医院管理信息系统(PACS)和远程医疗系统的发展应运而生。DICOM 的定制为不同的医学影像设备和用户提供了统一的接口标准和交互协议。DICOM 医学图像在 CT、MR、医院 PACS 系统中应用广泛。编程实现 DICOM 医学图像格式转换 ,对医学影像技术的研究具有重要的意义。

关键词 :DICOM ;BMP 格式转换

中图分类号 :TP311

文献标志码 :A

为了提升医疗服务水准 ,跟随医学影像技术和网络技术的发展 ,我国众多医院都在大量配备各种数字影像设备 ,并开始建立医院管理信息系统(HIS) ,开展远程医学试验。他们对于图像存档和传输系统(PACS)的需求越来越紧迫 ,PACS 的应用前景十分诱人。PACS 必需解决的技术问题之一是统一各种数字化影像设备的图像数据格式和数据传输标准。为此 ,诞生了新的医学数字成像及通信标准 ,即 DICOM (Digital Imaging and Communications in Medicine)。DICOM3.0 标准自发布以来 ,成为事实上的行业标准 ,各医疗设备生产厂商都严格按照此标准来制定自己的医学图像格式。生产的设备都带有 DICOM 格式的图像文件接口 ,可以方便地进行图像的传输和存储。对医学图像的后续处理必须先解读 DICOM 医学图像文件。课题来源是数字化虚拟人数据的图像处理 ,为了方便虚拟人的 CT 和 MRI 数据的图像配准 ,先把它们从 DICOM 格式转化为 BMP 格式。

1 DICOM 文件格式

1.1 数据元和数据元素的基本结构

DICOM 文件一般由 1 个 DICOM 文件头和 1 个 DICOM 数据集组成 ,如图 1 所示。

DICOM 文件头(DICOM File Meta Information)包含了标识数据集的相关信息^[1]。文件头的最开始

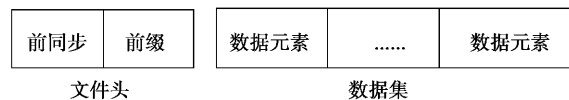


图 1 DICOM 文件总体结构

是文件前言 ,它由 128 个 00H 字节组成 ,接下来是 DICOM 前缀 ,它是一个长度为 4 字节的字符串 “ DICM ” ,可以根据该值来判断一个文件是不是 DICOM 文件。文件头中还包括其它一些非常有用的信息 ,如文件的传输格式、生成该文件的应用程序等等。

在 DICOM 文件中最基本的单元是数据元素(Data Element)。DICOM 数据集就是由 DICOM 数据元素按照一定的顺序排列组成的。DICOM 数据元素的组成如图 1 所示。它主要由 4 个部分组成 :标签、VR (Value Representation ,数据描述)、数据长度和数据域。

标签是一个 4 字节的无符号整数。DICOM 所有的数据元素都可以用标签来唯一表示。在 DICOM 中人为地将标签分为两个部分 :组号(高位 2 字节)和元素号(低位 2 字节)。表示值 VR 指明了该数据元素中的数据是哪种类型的。在 DICOM 文件中 ,它是一个长度为 2 的字符串。在数据元素中 ,VR 是可选的 ,它取决于协商的传输数据格式。DICOM 中规定了显式 (Explicit VR)和隐式(Implicit VR)两种传输格式 ,其

收稿日期 2007-06-07

作者简介 彭承琳(1936 -) ,男 ,重庆大学教授 ,博士生导师 ,主要从事医学信息检测与医学图像处理方向的研究 (Tel)023-65111954 (E-mail)pengchenglin@cqu.edu.cn.

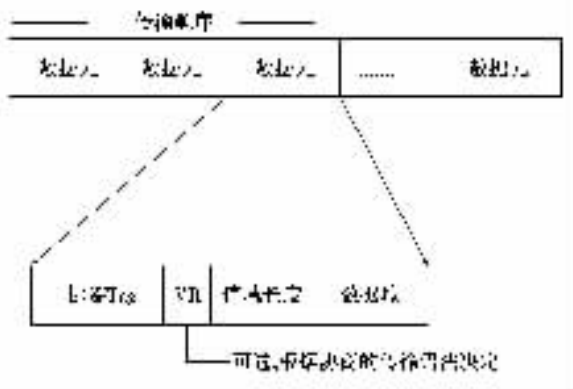


图 2 DICOM 数据元

中在显式传输时,VR 必须存在,在隐式传输时,VR 必须省略。

值域长度指明该数据元的数据域中数据的长度(字节数)。

数据域中包含了该数据元的数值。DICOM 中所有的数据都是以数据元素的形式出现的。

SQ 数据元和嵌套数据集的格式,表示值为 SQ 的数据元,其值域可以包含一个或多个数据元,这些数据元称为项。每个项中又可以包含其他项,从而形成多级嵌套结构。SQ 数据元中有 3 个特殊的数据元,即项(FFFE, E000),项定界数据元(FFFE, E00D),系列定界数据元(FFFE, E0DD)。不管传输语法所规定的表示值编码规则是什么,这 3 个数据元都必须采用隐含表示值编码方式。

1.2 传输语法

传输语法是由一个 UID 标识的。DICOM 默认的传输语法是隐式小端字节序(Implicit VR Little Endian)传输语法,并采用无损方式的 JPEG 压缩算法。

传输语法定义了 3 方面的内容:数值表示法如何指定;多字节数在存储或传输时的字节顺序,是低位字节(Little Endian)先存储发送,还是高位字节(Big Endian)先存储或发送;封装情况下的压缩格式,是采用 JPEG 还是 RLE 的压缩算法,是有损方式还是无损方式等。例如,对于一个 32 位无符号整数“12345678H”,在 Little Endian 方式下的字节顺序为 78、56、34、12,而在 Big Endian 方式下的字节顺序则为 12、34、56、78。

1.3 图像编码

像素数据元(7FE0, 0010)是 DICOM 文件中最重要 的一个数据单元,该数据单元包含了医学图像的显示所必须的数据^[2]。与像素数据元有密切关系的其他数据元是:

- (0028, 0008) 图像帧数
- (0028, 0010) 图像行数

- (0028, 0011) 图像列数
- (0028, 0100) 分配位数
- (0028, 0101) 存储位数
- (0028, 0102) 最高位数

像素数据的编码由分配、存储和最高位数决定,分配位数一定是大于存储位数的。像素数据可以是压缩的,也可以是未压缩的。以压缩格式来传送数据时,表示值 VR 为 OB;以未压缩格式传送数据时,表示值 VR 为 OW。对于未压缩像素的数据,通常按照从上到下,从左到右的顺序,作为连续的位流进行编码和存储。对于压缩的像素数据,可以分段存放,由系列定界项表示结束,从而支持长度预先未知的图像压缩过程。

1.4 DICOM 文件示例

上面分析了 DICOM 文件的结构和规范,对 DICOM 标准和文件有了一定的认识。为了更好的了解 DICOM 文件格式,通过 UltraEdit 软件打开一个 DICOM 文件,也可以直接将 DICOM 文件的后缀名直接改为 .cpp 效果一样。文件的十六进制字节序列如图 3 所示。可以看出,00000000H——0000007FH 的 128 个字节 00H 为文件头^[3]。接下来的 4 个字节,ASCII 字符为‘DICM’。从 00000084H 开始,是由一系列的数据元组合而成的数据集。第一个数据元的完整字节为 02 00 00 55 4C 04 00 46 00 00 00。由于字节顺序的影响,必要时应交换字节顺序。

02 00 00 00 为标签,交换字节后为 0002 0000,在数据字典里的定义为该组的字节数大小。

55 4C 为表示值 VR,ASCII 字符为 UL,表示无符号 32 位整数。

04 00 为值域长度,交换后为 00000046,表示数据元素的值为 70,也是从该数据元结束到下一个组 0008 之间的有 70 个字节。

DICOM 标准第 6 部分定义的数据字典包含了 DICOM 文件中所有标签的定义,各标签有具体的含义。标签(0002, 0010)定义了文件的传输语法,文件传输语法为 1.2.840.1008.1.2.1,Explicit Little Endian。传输语法决定了文件各数据元的编码方式。

像素数据元(7FE0, 0010)从 0000041CH 开始,表示值为 OW,根据规定,表示值为 OW 的数据元编码为:标签 4 字节;表示值 2 字节;保留字 2 字节;值域长度 4 字节;值域,根据值域长度的值确定。由文件可知,像素数据元的值域长度为 00080000H,即 524288 个字节。

关于 BMP 的文件格式,在这里就不做说明了。

```

00000000h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000010h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000020h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000030h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000040h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000050h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000060h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000070h: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 ; .....
00000080h: 44 49 43 4D 02 00 00 00 55 4C 04 00 46 00 00 00 ; DICM...UL..F...
00000090h: 02 00 01 00 4F 42 00 00 02 00 00 00 01 00 02 00 ; ...OB.....
000000a0h: 10 00 55 49 14 00 31 2E 32 2E 38 34 30 2E 31 30 ; ..UI..1.2.840.10
000000b0h: 30 30 38 2E 31 2E 32 2E 31 00 02 0D 12 00 55 49 ; 008.1.2.1.....UI
.....
00000400h: 30 32 34 2E 30 30 28 00 53 10 44 53 02 00 31 20 ; 024.00(.5.D5..1
00000410h: E0 7F 00 00 55 4C 04 00 0C 00 08 00 E0 7F 10 00 ; ?..UL.....?..
00000420h: 4F 57 00 00 00 00 08 00 44 00 44 00 44 00 44 00 ; 0W.....D.D.D.D.
00000430h: 45 00 45 00 44 00 42 00 40 00 40 00 41 00 42 00 ; Z.E.D.B.B.B.A.B.

```

图3 DICOM文件示例

2 DICOM 格式转换为 BMP 格式

2.1 格式转换的思路和流程

首先,定义一些存储图像参数的变量,并将变量初始化;其次打开一个 DICOM 文件按 Explicit Little Endian 读取数据元(0002,0010),确定后续数据元的传输语法;再按照传输语法格式要求,对 DICOM 文件的各相关数据元进行遍历,提取出 0002 组、0028 组和 7EF0 组有用的数据元的内容,分别存入先前定义的变量,关闭 DICOM 文件,再将像素数据元(7EF0,0010)里的 12 位或者 16 位的图像数据转换为 8 位灰度数据^[4],最后打开一空白文件,按照 BMP 格式的要求,将从 DICOM 文件里提取出的数据写入新文件,格式转换完毕,将在同目录文件夹下生成同名的 BMP 文件。

2.2 实际转换过程中需注意的问题

1)获取图像信息,实际上是对 DICOM 文件中各数据元的遍历。在遍历数据元时,有些数据元中包含的信息是和图像无关的,为提高程序遍历的速度,只读取有用的数据元。

2)读取数据元(0002,xxxx)时,该组的所有数据元的表示值 VR 都是显示表示的,即 Explicit VR Little Endian。由于是小端字节顺序,首先应交换字节顺序。还要特别注意数据元(0002,0010),该数据元的值确定了文件的传输语法^[5-6]。

3)像素数据元里的(7EF0,0010)里的图像数据一般都是 16 位或者 12 位的,需要通过调整窗宽窗位将原始数据转换成 8 位灰度数据。所谓的窗宽是指图像数据 displays 的范围,窗位是指图像数据 displays 的中心值。可以根据以下方程进行窗宽窗位调整:

$$y(x) = \begin{cases} 0 & \rho - w/2 > x \\ 255 & \rho + w/2 < x \\ [x - (\rho - w/2)] \times 255/w & \text{其它} \end{cases}$$

其中 y 表示显示的位图灰度值, x 表示图像的数据, w 表示窗宽, ρ 表示窗位。

2.3 结果显示

将 GE 的一幅 MR DICOM 脑部图像转换成了 BMP 图像,效果如图 4 所示。

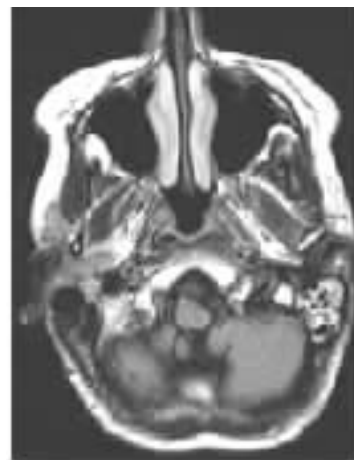


图4 转换后的 BMP 图像示例

3 结束语

DICOM3.0 作为医学图像存档和通信的国际标准,是所有医学影像技术的基础^[7]。由于大多设备厂商提供的 DICOM 图像显示软件都是基于 UNIX 系统的,因此将 DICOM 格式的图像转换成 Windows 平台下最常用最流行的 BMP 格式就很有实际意义,也为进行后续图像处理打下了基础。当然,在熟练掌握了 DICOM 文件格式后,将 DICOM 转换为其他格式也是

可以的,其原理和转换成 BMP 格式类似^[8-9]。通过这个格式转换软件,很方便地将虚拟人的 CT 和 MRI 数据格式转换到了 BMP,方便下一步编程实现 CT 和 MRI 的图像配准。

参考文献:

- [1] 林春漪,尹俊勋,马丽红,等. 论解读 DICOM 医学图像文件及其显示的 VC++ 实施方法[J]. 现代医学仪器与应用 2005,17(1):1-4.
- [2] 冯辉. DICOM 二进制文件解读的 VC++ 实现[J]. 医学影像杂志 2003,13(7):538.
- [3] 谢长生,熊华明. DICOM 图像显示的研究与实现[J]. 计算机工程与科学 2002,24(6):38-41.
- [4] 张尤赛,陈福民. DICOM 医学图像窗口变换的加速算法[J]. 计算机工程与应用 2003,(13):218-220.
- [5] PHILLI PV P. New medical imaging technology [J]. Biomedical Engineering, 1988(23):113-114.
- [6] LEVINE BA, CLEARY K R, NORTON G S, et al. Challenges encountered while Implementing a multi-center teleradiology network using DICOM3. 0. SPIE proceedings [C]//Medical Imaging 1997: PACS Design and Evaluation, Engineering and Clinical Issue 1997. 237-245.
- [7] 李刚荣,李桂祥,李晴辉,等. 浅谈 PACS 系统[J]. 医疗设备信息 2005,20(7):30-44.
- [8] 张连强,孙延泉. 核磁共振图像与 BMP 图像间的格式转换[J]. 中国医疗器械杂志 2002,26(3):215-217.
- [9] 钟国康. DICOM 图像在 PC 机上的显示编辑和演示方法研究[J]. 中国医学教育技术,2005,19(1):58-61.

Conversion of Medical Image based on DICOM with VC++

PENG Chen-lin CHEN Cheng CHEN Yuanyuan

(College of Bio-engineering, Chongqing University, Chongqing, 400030 China)

Abstract: DICOM3.0 is the international standard of medical images archiving and communication. With the popularization of cyber-medical equipment and the development of telemedicine system, DICOM3.0 emerges. It provides the interface standard and communication protocol for different medical imaging equipments and their users. Medical Image based on DICOM is widely used in CT, MR and PACS in the hospital. Converting the medical image based on DICOM with VC++ is very important for medical image technology study.

Key words: DICOM, BMP format conversion

(编辑 陈移峰)

(上接第 125 页)

system. The model has the following features: The blood circulatory system is treated as an independent system in the model; the interaction between the internal organs and the overall responses of circulatory system is emphasized. The descriptions of the characteristics of the various components are simplified so that fewer parameters and less calculation are required to develop the physical model with more explicit physiologic significance. A better approximation of the true physical testing system is realized since the new model simulates the paths of blood flow directly. As a result, the experiments have proved that the mathematic modeling results are corresponding with the test results.

Key words: total artificial heart test system; flow path mathematical model; elastic chamber model

(编辑 吕建斌)