Western Norway University of Applied Sciences

BACHELOR THESIS:

Development of a python based script for conversion of DICOM medical images to MCNP input files for Monte Carlo simulations

<ChengYi>

31. May. 2019

Document Control

Report title BO19E-65 Bachelor	*Date/Version* April 19th 31. May. 2019/0.13 *Report number:* BO19E-65
Author(s): ChengYi	*Course:* ELE150 *Number of pages including appendixes* 50
Supervisor at Western Norway University of Applied Sciences Professor Ilker Meric, Icaro Valgueiro malta Moreira	*Security classification:* Open
Comments: We, the authors, allow publishing of the report.	

Contracting entity: Western Norway University of Applied Sciences	*Contracting entity's reference:* Organization number: 917641404
Contact(s) at contracting entity, including contact information: *Phone number: (+47) 55 58 58 00* *E-mail: post@hvl.no*	

Revision	**Date**	**Status**	**Performed by**
0.11	16.05.19	First issue	ChengYi
0.12	20.05.19	Revised issue	ChengYi
0.13	26.05.19	added Appendixes and Corrected index and numbering	ChengYi

Preface

This report is my bachelor thesis as an exchange student in the Western Norway University of Applied Science. I would like to thank the Western Norway University of Applied Science and my home university, Tianjin University of Technology for providing me with this opportunity of exchange study. I have learned a lot and grown a lot in this exchange life. The completion of the bachelor thesis also means the end of half a year's exchange study.

The research work of this paper is completed under the guidance of my supervisor, Dr. Ilker Meric. During the half-year study in the Western Norway University of Applied Science, I was deeply impressed by Dr. Meric's rigorous working attitude and solid academic foundation. He helped me a lot in the topic selection, basic knowledge learning, mid-term programming and paper writing. In the process of completing this graduation project, I have learned a lot of useful knowledge and gained a lot of valuable experience, which is valuable for my future study and work. I would like to express my sincere gratitude to Professor Dr. Meric

I would like to express my gratitude to Professor Emil. Professor Emil is my teacher of Control Engineering, and he has helped me a lot in both professional courses and bachelor project. His help and advice have benefited me a lot.

I would like to thanks Icaro for his help. Icaro gave me lots of important guidance in my study and programming in this bachelor project. Without his help, I would not have completed this graduation design so smoothly. Icaro is my best teacher and best friend both in school and in life.

Summary

In this report, we design a simple, fast and effective method to convert DICOM files into MCNP input files. The format of MCNP input file is very complicated, especially the description of geometric parameters of related objects is very difficult. The main tool of the conversion methods is Python. Using Pydicom read DICOM file image pixel values of the matrix, judge the material represented by the cell through the relationship between the pixel value and the CT value, and generate the MCNP input file with the MCNP input format of grid element card, so as to display the DICOM image through the MCNP software. Since the general DICOM file size is relatively large, in the conversion process, we get a relatively small MCNP input file by compressing the DICOM image. In adittion, an algorithm combining pixels with the same CT value is designed to further reduce the size of MCNP input file.

1 Index

1 Introduction

1.1 Contracting entity

Western Norway University of Applied Sciences (HVL) is one of the largest higher education institutions in Norway. HVL was formed through a merger between Bergen University College, Sogn og Fjordane University College and Stord/Haugesund University College on January 1 2017. The total number of students at HVL is about 16000, and there are 1800 academic and administrative staff. Its main campus is in the Kronstad neighborhood of Bergen, Norway.[1]

1.2 Problem description

With the development of human anatomy and the radiation physics of tumor, the computer aided therapy based on medical sequence image is playing an increasingly important role. The DICOM (Digital Imaging and Communications in Medicine) standard is the fundamental standard in digital medical imaging and communicating. The implementation of DICOM standard, which greatly simplifies the realization of the exchange of medical image information, has developed into the international standard of medical image information [2].

MCNP is a large neutron, photon, and electronic transport simulation program developed by the Los Alamos lab in the United States, which uses the Monte Carlo method to calculate neutrons, photons and electrons and their coupling transport problems [3], which play an important role in data simulation and simulation calculations.

In this project, we want to use MCNP software to display DICOM file images. Before using the MCNP program, the user needs to write an input file called INP file by himself [3]. However, the input file structure of the MCNP program is complicated, in particular, it is difficult to describe the geometric parameters of related objects [4]. And the process is time-consuming and the probability of error is greater, and sometimes has to be remade or simplified. In addition, the MCNP program limits the number of geometric sizes of the input model [5], and if the number is beyond that limit, the MCNP program will be not working properly. However, the physical structure of the human body is extremely complex [5], and the complete and accurate surface of the MCNP program is difficult and sometimes impossible.

Therefore, the problem is how to automatically write MCNP input files from DICOM files

1.3 Main ide of solution

In this project, we chose to use python as the primary tool for converting DICOM files into MCNP input files. There is a library of instructions in the python library specifically for handling DICOM files called Pydicom. Pydicom is a pure Python package for working with DICOM files such as medical images, reports, and radiotherapy objects. Pydicom makes it easy to read these complex files into natural pythonic structural for easy manipulation [6]. Modified datasets can be written again to DICOM format files.

Using Pydicom read DICOM file image pixel values of the matrix, judge the material represented by the cell through the relationship between the pixel value and the CT value, and generate the MCNP input file with the MCNP input format of grid element card, so as to display the DICOM image through the MCNP software. Since the general DICOM file size is relatively large, in the conversion process, we get a relatively small MCNP input file by compressing the DICOM image. In addition, an algorithm combining pixels with the same CT value is designed to further reduce the size of MCNP input file.

2 Specification of MCNP input file

The MCNP program calculated by reading in a user-created input file called INP. The file must be organized according to the format of grid element card, and specify the information to describe spatial problems, specifically [5] :

(1) Description of spatial geometry

(2) Description of materials used in geometry and selection and estimation of cross regions;

(3) Description of the position and characteristics of neutron, photon and electron particle sources

(4) The types of necessary answer CARDS and mark CARDS;

(5) Any necessary redundancy elimination technology to improve the computational efficiency.

In this project, the first three parts are necessary, which are called cell card, data card and surface card respectively. The last two parts can be omitted for explanation. Therefore, the main problem of this project is how to write these three parts according to the DICOM file. The specific structure of these three parts is as follows:

1. Surface card: surface card is used to define the geometry of cell. Surfaces can be defined by equations, points, or microbodies. In this work we use microbodies to define the geometry, box. The surface number is the first entry, then is the x, y, x coordinates of a corner of box. After that is 3 vectors of 3 sides from the specific corner coordinates. Variables and their explanations are shown in table 2.1.

Table 2.1 The paramater of the “box“ microbody

Input Parameter	**Description**
Vx Vy Vz	The x,y,z coordinates of a corner of box
a1x a1y a1z	Vector of 1st side from the specified corner coordinates
a2x a2y a2z	Vector of 2nd side from the specified corner coordinates
d2x a2y a2z	Vector of 1rd side from the specified corner coordinates

For example, there is a row of surface card data that defines a square.

BOX -1 -1 -1 2 0 0 0 2 0 0 0 2

This input represents a cube centered at origin, 2 cm on a side, with sides parallel to the major axis.

2. Cell card: the cell number is the first entry and must begin in the first five columns. The next entry is the cell material number, which is arbitrarily assigned. The material is described on a material card

(data card) that has same material number. If the cell is void, a zero is entered for the material number. Next is material density. A positive entry is interpreted as atom density in units of 10^23 atoms/cm^3. A negative entry is interpreted as mass density in units of g/cm^3. In this work, it is all atoms density. The last one is the surface number; a negative entry is interpreted as inside the box.

This is an example of cell card:

```
c  cell card

1    1   -1.0   -1  2  -3  4  -5  6    $ 1 mm3 voxel filled with water (centered at 0,0,0)

2    0    -4   (-1:2:3)               $ Inner world filled with vacuum

3    0     4                          $ Outer world filled with vacuum
```

3. The data card follows the second blank card delimiter or third blank card if there is a message block. The card name is the first entry and must begin in the five columns. The required entries follow, separated by one or more blanks. Through various databases on the web, it is easy to get Atom Fractions and densities of common body tissues.[8][9][10]

This example is the data card of water:

```
c  Data cards

c  Water

m1   1000.    2.0  $H

     8000.    1.0  $O
```

3 Problem analysis

3.1 Description of the problems

1. As we know, in the end we should output a MCNP input file, which is in fact a standad form of a TXT file made up of element cards. So the first problem is how to write the corresponding TXT file in the prescribed form in Python as the MCNP input file.

2. DICOM image is composed of many pixels, but the pixel value of DICOM image is not the real CT value. How to determine the tissue contained in the location of the cell by connecting the relationship between the pixel value of image and the CT value is an important issue.

3. As the matter of fact, DICOM file image size is generally large. In this project, the example DICOM file we used is the head CT image of the human body, which image size is 1024*1024. However, such a high resolution is not required when the actual treatment plan is made, so we need to properly process it and reduce the resolution to improve the running speed of MCNP software. In addition, an image with 1024*1024 resolution has 1048,576 pixels, while 100 or so CT slices have more than 100 million pixels, and MCNP can theoretically withstand the maximum cell number of 99,999,999.Therefore, how to combine cells with the same material is the key issue of this topic.

3.2 Description of the solution

1. In this project, we use python as the main tool for converting files. Specifically, we use OS library for file operation, process DICOM files through Pydicom library, and extract related image information. Numpy library is used to process the pixel matrix of the extracted image, and finally write to TXT file according to the MCNP input file format.

2. After CT scanning, a global offset value is often added during data transmission or storage to make the pixel value of the DICOM file a positive integer, so in order to get the real CT value, the read DICOM value should be reversed. By searching data, we obtained this equation:

$$HU = Pixel\_value * slope + intercept$$

The CT value of the CT image reflects the X-ray absorption value of the tissue (attenuation coefficient u), and its unit is Hounsfield Unit (Hu), referring to the attenuation coefficient of water, i.e., the CT value of water is 0; Material failure The subtraction coefficient is positive if it is greater than water, and negative if it is less than water. The attenuation coefficients of bone cortex and air were taken as the parameters The limits and the lower limits, let's say plus 1,000 and minus 1,000.The rescale slope value and rescale intercept value read from DICOM file can be used to convert each pixel value into CT value, and then the tissue material corresponding to each pixel can be determined by judging the range of CT value. Mean and standard deviation (SD) of the HU values 7 tissue-equivalent phantom materials[9] are shown in table 3.1.

Table 3.1 7 tissue-equivalent phantom materials

Material	Mean(HU)	SD(HU)
Soft issue	24	9
Brain	52	8
Spinal disc	92	2
Trabecular bone	197	7
Cortical bone	923	107

Tooth dentin	1280	27
Tooth enamel	2310	80

3. In order to reduce the number of boxes that need to be defined, the program is mainly divided into the following two steps:

（1） Delete all pixel blocks with a pixel value of 26. Since the pixel value of 26 represents air, it is not necessary to display it on MCNP.

（2） Merge the adjacent pixel blocks belonging to the same material in the same row, and the result is cuboid. Each cuboid is defined as a box. The schematic diagram is shown in figure 3.1.

4 Realization of the solution

1. Firstly, we extracted the pixel matrix of DICOM image through pydicom. Since the matrix size was too large, we used numpy to sample every 8 pixels of the pixel matrix and compressed the resolution of 1024*1024 to 128*128.

2. Secondly, by designing a merging algorithm, adjacent pixels belonging to the same material in the same row are merged into a cuboid. The algorithm steps are as follows:

(1). Start from the first pixel in the first row, find the first pixel whose pixel value is bigger than 26, and record its spatial position as the first pixel of the cuboid.

(2). Search the next pixel sequentially. If the pixel is equal to the value of the previous pixel, add 1 to the length of the cuboid; If the pixel is not equal to the previous pixel, end the current pixel merge and output the position, length, pixel value and other information of the previous cuboid. Loop step 2 all the way to the end of the line and go to the next line starting with the step.

The specific program flow chart is shown in figure 4.1.

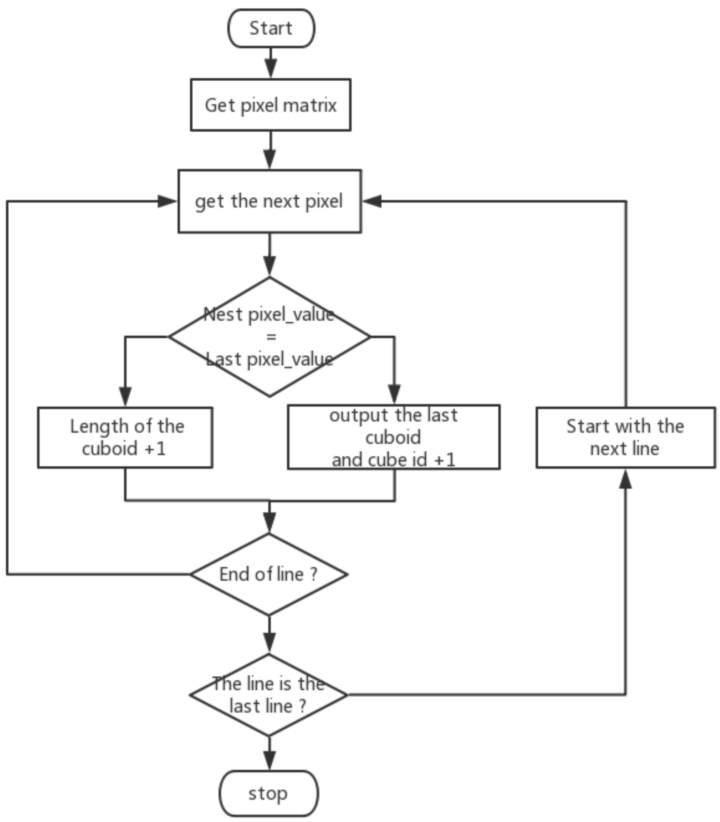

Figure 4.1 program flow chart

And by using this algorithm, we can get the surface card. The specific program of the surface card is shown in figure 4.2:

```
def surface_card():

    global cube
    pixel_array_down = ds.pixel_array[::16,::16]
    pixel_size = 0.1875*8
    sl = ds.get('SliceLocation')
    b = 0
    y = sl - 0.625
    lv = 26

    width = len(pixel_array_down)
    depth = len(pixel_array_down[0])
    for i in range(width):
        lv = 26
        b = 0
        for j in range(depth):
            v = pixel_array_down[i][j]
            if v != lv:
                if lv > 26:
                    cube = cube+1
                    xb = b*pixel_size
                    fp.write('\n'+str(cube)+' box
                if v > 26:
                    x = i*pixel_size
                    z = j*pixel_size

                b = 1
                lv = v
            else:
                b = b+1
                lv = v
```

Figure 4.2 the specific program of the surface card

3. Cell card: The slope and the intercept value read from DICOM file can be used to convert each pixel value into CT value, and then the tissue material corresponding to each pixel can be determined by judging range of CT value. Then write the cell card in the format. The program flow chart of the cell card is shown in figure 4.3.

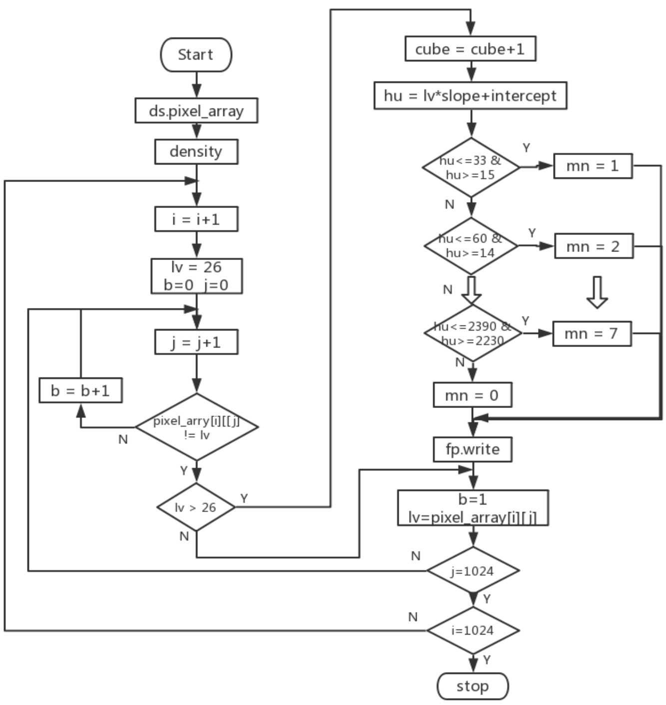

figure 4.3 the flow chart of cell card

The specific program of the cell card is shown in figure 4.4:

```
def cell_card():
    global cube
    pixel_array_down = ds.pixel_array[::16,::16]
    slope = ds.RescaleSlope
    intercept = ds.RescaleIntercept
    mn = 0
    density = {"1":"0.099006","2":"0.104021","3":"0","4":"0.133974","
    width = len(pixel_array_down)
    depth = len(pixel_array_down[0])
    lv = 26
    b = 0

    for i in range(width):
        lv = 26
        b = 0
        for j in range(depth):
            v = pixel_array_down[i][j]
            if v != lv:
                if lv > 26:
                    cube = cube+1
                    hu = lv * slope + intercept
                    if hu <= -998:
                        continue
                    elif hu <=33 and hu >=15:
                        mn = 1
                    elif  hu <=60 and hu >=44:
                        mn = 2
                    elif hu <=94 and hu >=90:
                        mn = 3
                    elif hu <=204 and hu >=190:
                        mn = 4
                    elif hu <=1030 and hu >=816:
                        mn = 5
                    elif hu <=1307 and hu >=1253:
                        mn = 6
                    elif hu <=2390 and hu >=2230:
                        mn = 7
                    else:
                        mn = 0

                    fp.write('\n'+str(cube)+'    '+str(mn)+'    '+str(c

                b = 1
                lv = v
            else:
                b = b+1
```

figure 4.4 the program of cell card

5 Testing

In this case, we used Moritz software to display the MCNP 3-dimensional images.[10]

Take the 80_dicom_125mm.dcm file as an example. First draw it using the matplotlib.pyplot instruction in python as shown in the figure 5.1.

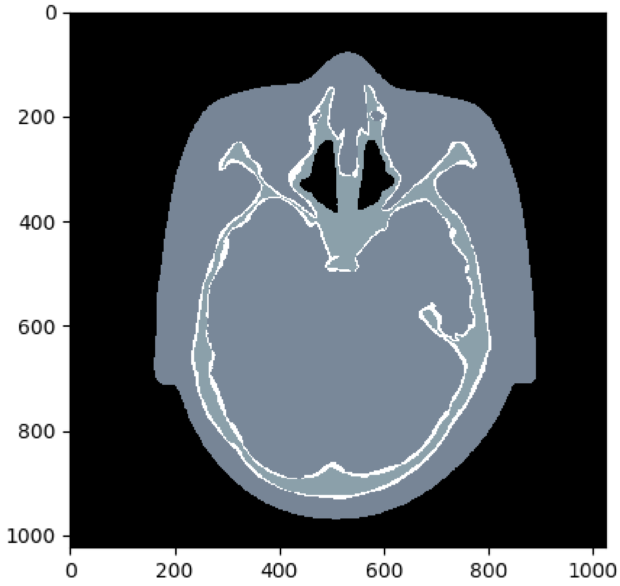

figure 5.1. 2D slice through the head phantom generated using python

Convert DICOM file into MCNP input file through the program designed by us. See the appendix for this report. Then run the input file with MCNP software, and the output figure is shown in the figure5.2. In this situation the image size is 1024*1024.

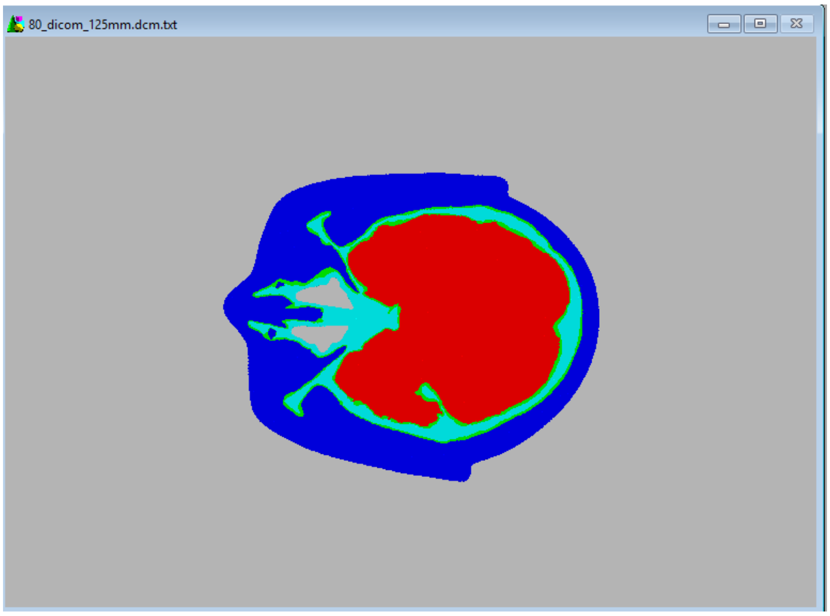

figure 5.1. plot using MCNP (1024*1024)

After further compression, the resulting image of slience 80 is shown in figure 5.2.

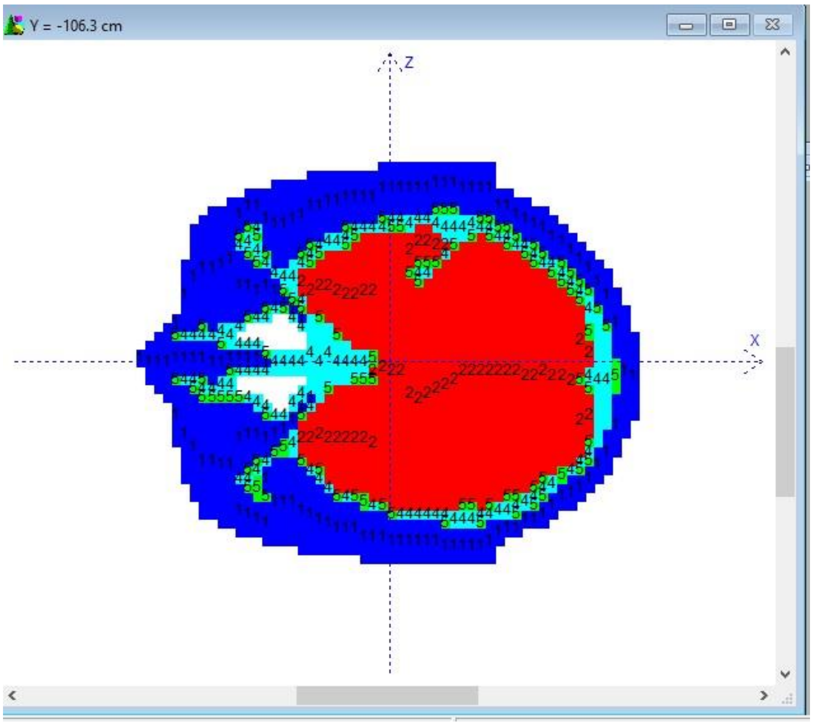

figure 5.2. plot using MCNP (128*128)

By combining images from different slices, we get a three-dimensional image like the figure5.3.

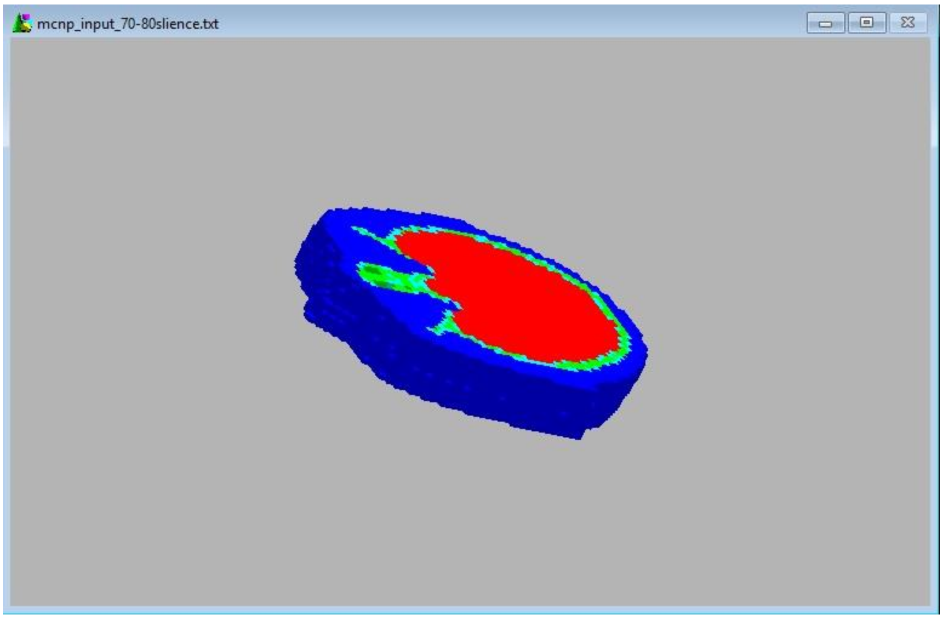

figure5.3.1 slice 70-80

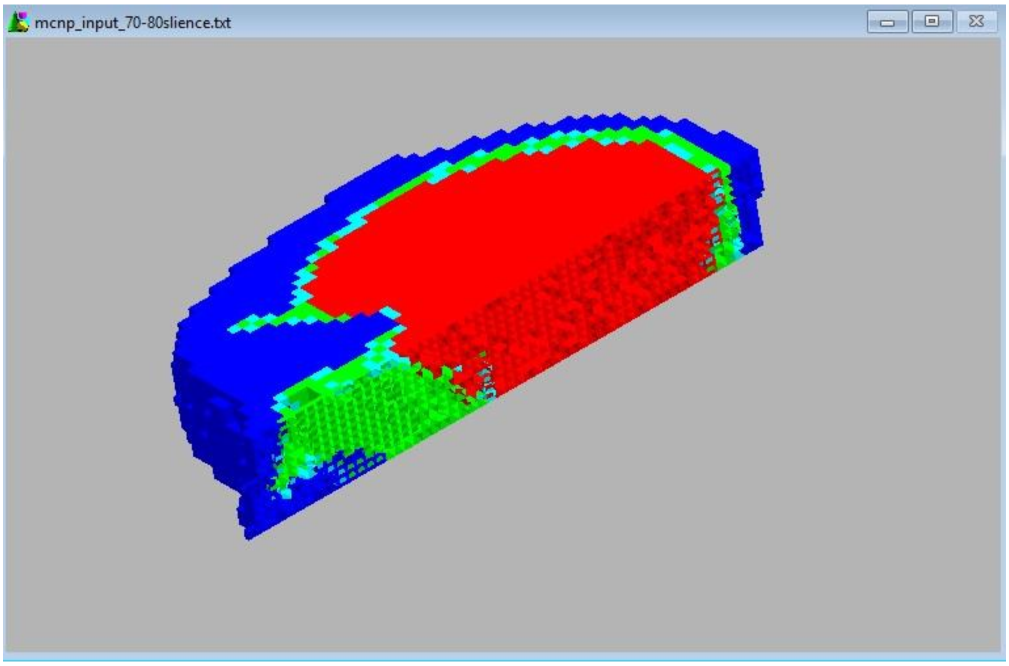

figure5.3.1 slice 70-80_Zclip

6 Conclusion

In this project, we mainly solved the problem of how to convert DICOM files into MCNP input files. In order to reduce the size of the MCNP input file, we reduced more than 1 million rows of data to less than 500 rows by compressing the resolution and merging the pixels. The MCNP running speed is greatly increased. However, there are about 100 CT images of a person's head, and it can take more than an hour to display them all at once. Therefore, in future work, we also need to design better algorithms for merging pixels to further reduce the file size.

Although we have designed a program that can successfully convert DICOM files to MCNP input files, it has poor operability. If people want to use this program to generate DICOM files, they need to copy DICOM files to the specified file directory, or change the file directory in the program, which is relatively troublesome. In future work, we should develop a visualization script for MCNP input files based on the existing program. This will greatly enhance its usefulness.

Reference

[1] Western Norway University of Applied Sciences . https://www.hvl.no/en/about/

[2] Mildenberger P , Eichelberg M , Martin E . Introduction to the DICOM standard[J]. European Radiology, 2002, 12(4):920-927.

[3] Briesmeister J F, (2000) MCNP–A general Monte Carlo N-particle transport code. Los Alamos National Laboratory LA-13709-M

[4] 陈卓, 刘晓平, 施灿辉. 基于 MCNP 的医学仿真计算建模方法研究[J]. 系统仿真报,2004(10):35-38.

[5] Pelowitz, Denise B. "MCNP6 user's manual version 1.0.", LA-CP-13-00634, Rev. 0, (2013).

[6] X G Xu, T C Chao, A Bozkurt. Vip-Man: An Image-Based WholeBody Adult Male Model Constructed From Color Photographs Of TheVisible Human Project For Multi-Particle Monte Carlo Calculations [J]. Health Physics, 2000, 78(5): 476-486.

[7] Getting Started with pydicom .https://pydicom.github.io/pydicom/stable/getting_started.html

[8] Yoriyaz H., Santos A., Stabin M. and Cabezas R., Absorbed fractions in a voxel-based phantom calculated Indexwith MCNP-4B code. Med Phys. vol 27, p 1555-1562, 2000.

[9] McConn, Ronald J., Gesh, Christopher J., Pagh, Richard T., Rucker, Robert A., and Williams III, Robert. *Compendium of Material Composition Data for Radiation Transport Modeling*. United States: N. p., 2011. Web. doi:10.2172/1023125.

[10] Giacometti, V., Guatelli, S., Bazalova-Carter, M., Rosenfeld, A. B. & Schulte, R.W. (2017).Development of a high resolution voxelised head phantom for medical physics applications. Physica Medica: an international journal devoted to the applications of physics to medicine and biology, 33 182-188.

[11] Defining MCNP Sources in Moritz. (2009). Nuclear Technology, 168(3), 848–851. https://doi.org/10.13182/NT09-A9317