

# A Novel Scheme in Lossless Compression of Medical Image

Hongying Han & Yanguang Liu & Qiuling Wang & Yan Li & Tengfei Li & Jianfeng Chu\*

*Jilin University, Changchun, 130012, China*

Haiou Huang

*Jilin Agriculture Science and Technology College, Jilin, 132101, China*

*\*Corresponding author: Jianfeng Chu*

**ABSTRACT:** With the arrival of digital era as well as networked era, it is inevitable for medical field to realize digitization and network. Nowadays medical imaging has been in widespread use. However, the increasing number of images and their large-size character put much pressure on the storage space. What's more, the above factors are burdensome on network. In addition, they also burden the system when many doctors access to the system, which results in congestion. All of those have a negative effect on hospitals' efficiency. Thus, there is an urgent need to solve the above mentioned problems fundamentally. We aim to save time on the wire and save the storage space by increasing the calculation cost of the terminals. According to the analysis of the experimental data, this algorithm is capable of compressing the medical images which are in line with DICOM standard in a lossless form. Furthermore, this method has helped us get a better compression ratio. We compress 200 medical images realistically at the compression ratio of 3.85. In this way, the scheme can reduce the consumption of the network bandwidth and improve the transmission rate nearly 4 times, which saves time on network. It just increases the cost of the computing on the terminal.

**KEYWORD:** the image processing; the lossless compression; the boundary extraction

## 1 INTRODUCTION

With the advancement of medical treatment, CT (Computed Tomography Image), MR (Magnetic Resonance Image) and other digital medical imaging are widely used in medical diagnosis, which has also already moved towards digitization (Bairagi V.K et al, 2012). In order to promote the Digital Image Information's compatibility among devices of different manufacturers, set up the communication standard of the same and different areas of diagnostic equipment, build a unified information database, and to expand the picture archiving as well as communication system which interacts with the hospital information system. The American College of Radiology (ACR) and the National Electrical Manufacturers Association (NEMA) have set up a joint committee called ACR-NEMA in 1983. The joint committee set two standards (ACR-NEMA1.0 and ACR-NEMA2.0) in 1985 and 1988 respectively. But the two standards were not recognized widely because the technology was not mature at that time. Nonetheless, many instrument manufactures have joined ACR-NEMA in these years (Bairagi Vinayak K et al, 2013). After several years of hard work, the joint committee has set international standards of medical imaging and related information, namely the DI-

COM (Digital Imaging and Communications in Medicine) Standard. With the development of technology, nowadays the Dicom3.0 has grown into an international general standard in the field of medical images (Dos Santos DP et al, 2013) (Mahmoud Ismail et al, 2013). The statistic shows that several ten billions of medical images which approve of DICOM standard are used in nowadays' clinical application.

Nowadays, there is no exception that hospital deals with mounts of medical images every day. However, the size of medical image is more often than not larger than the common image. For example, the size of the Computed Radiography Image (CR) is approximately between 3000KB and 12000KB. However, the MB bandwidth network is capable of transferring merely around 10000KB of data, (So-called MB bandwidth network means definitely not 100000KB per second, actually it transmit data at a speed of 10000KB per second.) which puts much pressure on network transmission, and results in ca ton phenomenon when more than one doctor access to the medical image query system at the same time. We have to admit that it is hardly accessible to solve this problem merely through increasing the computer processing speed or enhancing the width of network channel. Therefore, it is essential

for finding the solution to compress the medical images.

## 2 PRELIMINARIES

### 2.1 The Composing of DICOM

As for DICOM files, what is in strictly accordance with the DICOM standard is medical file storage, suffix for DCM. File. Except for including the image data, it also includes patients' information, image technology parameters, and the doctors' diagnosis information, etc. Begin with 128 bytes (file header), common file structure is used to maintain compatibility. There is no special meaning in hexadecimal 00 as filler. Then four bytes of the string "DICM" follows. The part behind the file information is managed in the form of data set. According to certain standards, DICOM data set can be divided into several groups. Each group has a different group number, 0002 stands for the group of equipment communication, 0008 represents the characteristic parameter group, 0010 is the group of patients' information, 0028 is the group of the parameter set and image information, and 7FE0 represents the group of image data sets. Data set is composed of many data elements. The composition of the data element is shown in fig 1.

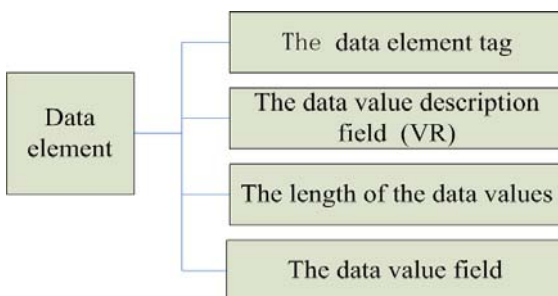


Figure 1. The composition of the data element

### 2.2 The Existing Boundary Extraction Algorithms

Image border extraction is a vital part of computer image processing, which has a wide application in various fields. The boundary of the existing extraction technology has the following several types, one of which is edge detection: Using the image of a derivative or the second derivative of zero information to extract the edge. Commonly used edge operators for convolution operation are Sobel operator, Roberts's operator, and Canny operator. They give detection of sharp changes among adjacent points in the image, finding the closed curve which is exactly the image boundary. The disadvantage of this method is that it is difficult to deal with the contradiction between the noise resistance and detection accuracy. If we improve the detection accuracy, we will be misled by the noise and take it as a nonexistent edge,

however, if we improve the noise resistance, it will cause inaccurate edge detection. Another existing extraction technology is the boundary integral method, including the contour extraction and boundary tracking. They have plenty of differences, for instance, contour extraction is aimed at looking for internal boundary points, then we take them as the background color, and the rest is regarded as the border. When it comes to boundary tracking, we always start from a boundary point. According to certain detection rules, the boundary tracking won't stop finding out other pixels on the target object contour until it returns to the starting point. What's more, as is mentioned above, the boundary extraction has to do two-valued processing upon images in pretreatment, which results in unavoidable damage to the image.

## 3 A NOVEL SCHEME IN LOSSLESS COMPRESSION OF MEDICAL IMAGE

The main scheme includes the following several parts. First of all, we extract the boundary. Secondly, we extract the graphics of the boundary. When faced with many graphics of complicated relations, we deal with the relations among the graphics in a hierarchical way. And then we store the boundary with the method of bit operation. Finally, we process the boundary layer-by-layer according to the above method.

### 3.1 The Boundary Extraction

On the of the basic operations of the medical images processing is the boundary extraction, which is based on the different pixel threshold of the boundary region and the background region.

When we begin the first scanning and find out the first point whose pixel is different from the background of the image, we can find the points which belong to the boundary extraction from its eight neighborhood points as shown in fig 1. Then we center on the point which is confirmed just now and judge whether the points which is from its eight neighborhood points agree with the principle of eight adjacent above or not. Loop until we find the point whose eight neighborhoods don't meet the principle of eight neighborhoods above. Following the above steps, we are capable of seeking other eligible pixels.

When the boundary point in its eight adjacent domains has been found, we continue to search for the eight neighborhood points centered on the boundary point and judge whether they accord with the three conditions of boundary point listed above or not. The following is an instance to account for the extraction procedure. As shown in fig 3, I ~VIII

show the sequence of scan, and 0~7 show the point relative to the previous point position.

### 3.2 The Graphics Extraction

After the whole boundary is extracted, we'll take the advantages of the information of the boundary points to dispose the inside ones. In most instances, the border is not a straight single line, but a gradual process in the actual medical imaging. So we need the judgment of the gradual process, it's simplified for late fill on pixel inside boundary. Processing method is listed, as follows:

1. Find out the extracted boundary points which have the same abscissa, and count the number of the boundary points.
2. Sort the boundary points with the same abscissa by the order from left to right.

### 3.3 The Compression in Form of Bit Operation

The latter point is competent of storing the location relative to the former point after storing boundary points' initial position accurately. The serial number of the relative location is stored in the form of bit operation. On account of only eight relative positions, each relative position is supposed to be shown by three binary bits. In this way, three bytes can store eight relative location points, which is capable of compressing the image.

## 4 EXPERIMENT

The program of the medical image compression is written by C, which can run in Windows 7 and Windows 8 system. The following picture is the example of the medical image.



Figure 2. The example of the medical image

200 medical images are used in our experiment (including CR, CT, DX, MR 50 each). We will compress the 200 images respectively by our new algorithm for lossless compression mentioned above. Results are as follows:

Using our new algorithm, the average compression ratio of CT is 3.64, the average compression ratio of CR is 3.74, the average compression ratio of DX is 4.19, the average compression ratio of MR is 3.83. So we can gain the average compression ratio of the 200 images: 3.85.

The following form is some experimental data:

Table 1. The Representative Experimental Data

Image type	The original size	The compressed size	The ratio of compression
CR	4478KB	1195KB	3.75
CR	5056KB	1518KB	3.33
CT	516KB	135KB	3.82
CT	623KB	169KB	3.68
DX	3688KB	903KB	4.08
DX	3697KB	984KB	3.76
MR	873KB	206KB	4.24
MR	525KB	127KB	4.13

Aiming at the MB broadband network, we may also work out the increment of transmission rate after compression. Supposing the network flow rate is 10000KB Bytes per second, the average size of the 200 medical images used in the experiment is 3039.96KB. Thus, it is easy to work out that 3.29 images can be transmitted per second. After compressed by our new algorithm at the compression ratio of 3.85, the average size of the 200 medical images is 789.6KB. Under such circumstance, images can be transmitted 12.66 per second.

There are many factors which exert a conductive influence on image processing. First and most, memory space is limited, thus compression attaches great importance to image processing. In addition, distortion is supposed to be avoided, spontaneously, lossless compression plays a vital role in image processing.

## 5 CONCLUSION AND FUTURE WORK

This paper provides a new compression method of the medical image, which presents a novel scheme about the boundary extraction. The scheme is aimed to make full use of the characteristics of medical image and increase the compression ratio as much as in possible.

Our planning for future work consists of two aspects. Firstly, we will spare no efforts on improving the compression ratio of our new algorithm. Our new algorithm gets the best compression ratio under an ideal situation which means there is no miscellaneous point. However, there are mounts of miscellaneous points which are other than boundary points in medical images actually. If we store them independently, it will waste much memory space. So we store their relative positions and relative color values. Secondly, we will focus on designing the

matching decompression process. According to the particular application of medical images, the compression of them must be lossless. Lossless compression means we are capable of restoring the original images free of error, namely, we utilize the compression images only in the process of transmitting and storing. What doctors use to diagnose is entirely accurate. After decompression, we must be sure to maintain the image and reduce the calculation cost of decompression as much as possible.

Apart from the image compression, the algorithm of the boundary extraction can also be used for the reconstruction of three-dimensional image. The reconstruction of the three-dimensional equals to get the structure information from the two-dimensional image, then restructure the three-dimensional model according to the tomography information of a series of the contours in the two-dimensional image. To realize the reconstruction of three-dimensional medical image, we segment and extract the input (the sequence of cross-section). Then describe the structure of the scene through the surface which consistent with element geometrical shape. Besides the extraction algorithm is capable of getting the structure information of the 2D image. The stereo-model may play an essential role in morphometry, anatomical orientation and operation design and simulation.

In the future, we may set up a digital information storehouse of human body through 3D imaging technology which based on boundary extraction. In the words, we need to extract the boundaries of one person's several medical images. What we get from the above practice ought to be dealt with 3D imaging technology. After integration, we are able to create on information group of the person's skeleton and organs. In this way, a digital information storehouse of human body can be created.

## REFERENCES

- A.P. Bhagat, "Medical image: formats, compression techniques and DICOM image retrieval a survey", Department of Computer Science and Engineering, Prof Ram Meghe College of Engineering and Management, Badnera, Amravati – 444701.
- Bairagi Vinayak K and Sapkal Ashok M, "ROI-based DICOM image compression for telemedicine", *Sadhana - Academy Proceedings in Engineering Sciences*, vol.38, no.1, pp.123-131, 2013.
- Bairagi V.K and Sapkal A.M, Automated region-based hybrid compression for digital imaging and communications in medicine magnetic resonance imaging images for telemedicine applications, *IET Science, Measurement and Technology*, vol.6, no.4, pp.247-253, July 2012.
- Dos Santos DP, Jungmann F, Friese C, Duber C and Mildemberger P, "Irreversible image compression in radiology. Current status", *Radiologie*, vol.53, no.3, pp.257-260, Mar 2013.
- Mahmoud Ismail and James Philbin, Multi-series DICOM: an extension of DICOM that stores a whole study in a single object. *J Digit Imaging* (2013) 26:691–697.