



HAL
open science

Segmentation of Small Animal Computed Tomography Images Using Original CT Values and Converted Grayscale Values

Guoqiang Ma, Naixiang Li, Xiaojuan Wang

► **To cite this version:**

Guoqiang Ma, Naixiang Li, Xiaojuan Wang. Segmentation of Small Animal Computed Tomography Images Using Original CT Values and Converted Grayscale Values. 7th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Sep 2013, Beijing, China. pp.470-477, <10.1007/978-3-642-54344-9_54>. <hal-01220970>

HAL Id: hal-01220970

<https://inria.hal.science/hal-01220970v1>

Submitted on 27 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Segmentation of Small Animal Computed Tomography Images using Original CT Values and Converted Grayscale Values

Guoqiang Ma^{1,a}, Naixiang Li^{1,b}, Xiaojuan Wang^{2,c}

¹ Dept. of Computer Science & Information Engineering, Tianjin Agriculture University, Tianjin 300384, China; ² Library, Tianjin Agriculture University, Tianjin 300384, China
^amgqwxj@163.com, ^blinaixiang@tjau.edu.cn, ^cwangxiaojuan6996@163.com

Abstract. Medical image segmentation is the foundation of normal and diseased tissue 3D visualization, operation simulation and visual operation. In this paper, we comprehensively use two values that represent the same object (body tissue), to segment by the same algorithm implementation. The original CT images are downloaded from the web source, the dog rib tumor CT scan image by GE medical systems, all the experiment of dataset of 312 thoracic CT scans. The core of the segmentation is k-means clustering algorithm. The segmentation process consisted of two phases: (1) convert CT value to JPG gray value or not use the original CT value as the data sets for clustering; (2) segmentation bone tissue using the new k-means clustering algorithm program which is implemented with MATLAB 2012a programming language and for two-dimensional data matrix directly. The experiment produced strikingly different results. These results may be indicating that not only the segmentation algorithm of CT image is important, but also the data for segmentation is important too.

Keywords: Image segmentation; Computed tomography; k-means clustering

1 Introduction

Image segmentation is a challenging task in image analysis. Image segmentation plays important role in medical image. A large variety of methods of have been proposed in several years. Intracranial image segmentation techniques for CT images include semi automated region-growing ^[1], intensity threshold ^[2, 3], and automated level set methods ^[4, 5].

The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. Clustering algorithm is a kind of do not need training, unsupervised (unsupervised) statistical method. In a sense, clustering is a classification of self training. Because of the lack of training sample set, so the iterative clustering algorithm is performed on the image classification and extraction of feature values. Among them, K-means ^[6], Fuzzy C-Means ^[7, 8], EM (Expectation-Maximization) and hierarchical clustering method ^[9, 10] is commonly used in clustering algorithms. K-means clustering

algorithm calculate the current average of each class, then according to the new generation of mean pixel to reclassify (mean pixel in recent class), the new generation class iterative execution front steps.

In this paper, we propose a new processing framework draws on the existing techniques of k-means clustering algorithm and use the CT value for segmentation in CT images. The main algorithm is still using the K-means algorithm. The best contribution of this paper are: (1) completed a specialized MATLAB language program implementation of k-means clustering algorithm for CT images segmentation; (2) using the CT value of DICOM format files directly to calculate by k-means clustering algorithm for segmentation. The results show the segmentation quality is better than using the transformational grayscale values.

2 DICOM format and original CT values

2.1 DICOM format

DICOM (Digital Imaging and Communications in Medicine) is a standard for handling, storing, printing, and transmitting information in medical imaging. It includes a file format definition and a network communications protocol. A DICOM data object consists of a number of attributes, including items such as name, ID, etc., and also one special attribute containing the image pixel data (i.e. logically, the main object has no "header" as such: merely a list of attributes, including the pixel data). A single DICOM object can have only one attribute containing pixel data. For many modalities, this corresponds to a single image. But note that the attribute may contain multiple "frames", allowing storage of cine loops or other multi-frame data. The same basic format is used for all applications, including network and file usage, but when written to a file, usually a true "header" (containing copies of a few key attributes and details of the application which wrote it) is added.

2.2 CT value

The characteristic of CT is able to distinguish the slight difference of human tissue densities; the standard is based on the body tissues of X-ray absorption coefficient (μ value) to decide. In order to calculate and convenience, the linear attenuation coefficient is divided into 2000 units, known as the CT value, unit of measurement called Hounsfield.

Using water as a value of 0, the upper bound of bone CT value is 1000; the lower air CT value is -1000. At present, most of the CT scanner has a range of 1000 or more than 2000. In fact, CT value is the corresponding X-ray value of the attenuation coefficient with the different human tissue in CT image. In the CT image, bone and air CT values were specified for upper and lower bounds. The calculation formula of CT value is shown in equation 1.

$$CT_Value(HU) = \frac{\mu_{tissue} - \mu_{water}}{\mu_{water}} \times 1000 \quad (1)$$

The CT values of normal human tissues, organs are shown in Table 1.

Table 1. CT values of normal human tissues

Classes	CT values(HU)
Water	0± 10
cerebrospinal fluid	3~8
plasma	3~14
edema	7~7
alba	25~32
ectocinerea	30~40
blood	13~32
clot	64~84
liver	50~70
spleen	50~65
pancreas	45~55
kidney	40~50
muscle	40~80
gallbladder	10~30
fat	-20~-80
calcification	80~300
atmosphere	<-200
bone	>+400

Matlab R2012a provides `dicominfo()` and `dicomread()` function, respectively, used to read DICOM format image file header data and metadata. For example, in the MATLAB R2012a command window input: `info=dicominfo('D:\ANONYMIZE2\DICOM7_000000.dcm')`, Running result will display detailed information about the file. Two parameters of these detailed information illustrate the transform relation of the record value and the real CT value, they are "RescaleIntercept" and "RescaleSlope". The conversional relation can be represented in equation 2.

$$CT_Value(HU) = pixel_value \times RescaleSlope + RescaleIntercept \quad (2)$$

CT_Value represents the real CT value of body tissue, *pixel_value* represents the pixel value of a DICOM formatted CT image file, it can be read out by the `dicomread()` function.

2.3 DICOM values covered to Grayscale values

Jpg/Jpeg is an efficient compression format, the file format is JPEG (Joint Photographic Experts Group) standard products. The depth image are 8bit, 12bit, considering the human eye to gray level of sensory discrimination, the general use of the bit depth is 8bit, the 256 gray level. In the gray image in JPG format, each pixel value is the gray value, the range between 0~255.

Conversion principle is very simple, is to use the `dicomread()` function to read the CT value matrix from a DICOM formatted CT image file and converted to a numerical matrix of 0~255. The conversional relation can be represented in equation 3.

$$I_{gray} = \frac{(I_{CT} - \min(I_{CT})) \times 255}{\max(I_{CT}) - \min(I_{CT})} \quad (3)$$

I_{gray} represents the 256 gray level image matrix after conversion, I_{CT} represents the CT value matrix read by *dicomread()* function.

Matlab key codes of the conversion program are as follows:

```
I=dicomread(CT_file);
I_ct=double(I);
    a=max(I_ct(:));
    b=min(I_ct(:));
    [x,y]=size(I_ct);
    for j=1:y
        for i=1:x
            I_gray(i,j)=(I_ct(i,j)-b)*255/(a-b);
        end
    end
imwrite(uint8(I_gray),jpg_filename,'jpg');
```

2.4 Matlab K-means clustering algorithm program

K-means algorithm is a hard clustering algorithm, it is a typical of the objective function clustering method based on the prototype of representative which uses the distance summation of data points to the prototype (class center) as the optimum objective function, using the method for the extremism of function to get the iterative adjustment rules. The K-means algorithm uses the Euclidean distance as the similarity measure, it calculates an optimal classification of vector $V = (v_1, v_2, \dots, v_k)^T$ based on an initial clustering center, and keep the evaluation value J_c to a minimum. Algorithm often uses the error sum of squares function as the clustering criterion function, the squared error function is defined as $J_c = \sum_{i=1}^k \sum_{p \in C_i} \|p - M_i\|^2$. The parameter M_i represents the average of class C_i , the parameter p represents a point of class C_i .

We do not use the k-means function in the MATLAB statistics toolbox and write a new k-means clustering algorithm program specifically for CT image.

The process of k-means clustering algorithm is improved by iterative method. The input parameters are two, respectively is the JPG format image file name $I_filename$ and the number of clusters K ; output only corresponding class label matrix IDX_mar . The termination condition function program is that the objective function value of error square sum is no change. The actual amounts of code of this process are only 38 lines using MATLAB programming language.

The pseudo - code of the improved k-means function as follows.

```
1 Begin:I_filename='jpg_file_name';k='num_class';
2 I=imread(I_filename);
3 [m n]=size(I);
4 Flag=zeros(m,n);
5 ma=max(I);mi=min(I);
```

```

6 Select k values as the center of initial K class from the mi to ma equidistantly
7 J=0;
8 For i=1 to m*n
9     For j=1 to k
10         Calculate the Euclidean distance from each point of I to K values of
class center
11         Save the minimum of K distance
12         make a kind of markers in Flag matrix
13     Next j
14 Next i
15 Calculate the error sum of squares
16 If Jc==J
17     Results: marked matrix Flag
18 End function: End
19 Else
20     J=Jc
21     According to the categories, recalculate the K centroid
22 Endif
23 Goto 8

```

The 256 gray levels JPG format image is segmented by the new k-means clustering algorithm program. This process is shown in Figure 1.

3 Materials and Method

The original CT images are downloaded from the web source, "DICOM7 medical image database", URL: <http://www.dicom7.com/>. The compression package is named for the "ANONYMIZE2", the dog rib tumor CT scan image by GE medical systems, all the experiment of dataset of 312 thoracic CT scans. The first step is converting the DICOM format images to JPG format images using the equation 1. The second step is segment the DICOM format images using CT values and the JPG images using gray values by the new k-means function.

The configuration of the computer used for experiments is with Intel Core Duo 3.33 GHz CPU, RAM Memory 2 GB, 32 bits Windows 7 ultimate; the algorithms are used MatLab2012a programming. We analysis the quality of segmentation results through observe segmentation result with the naked eye.

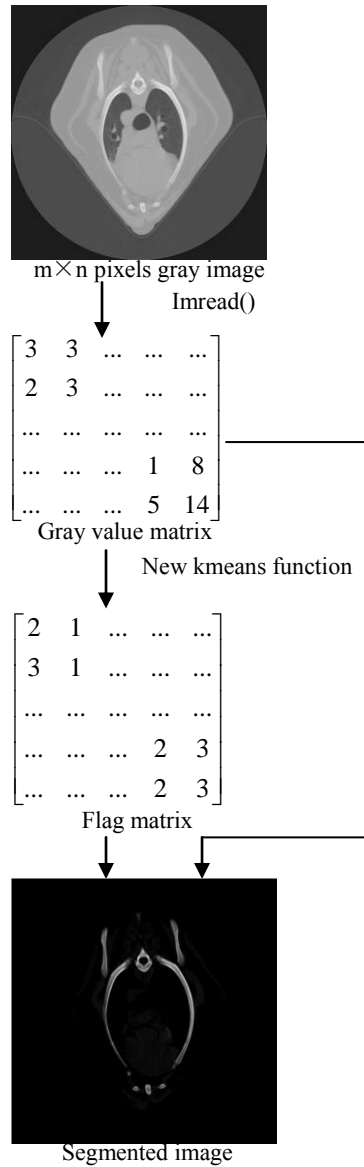



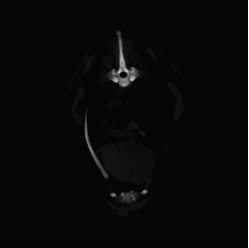
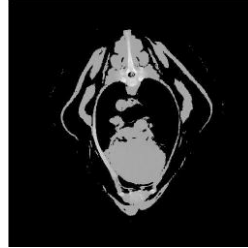
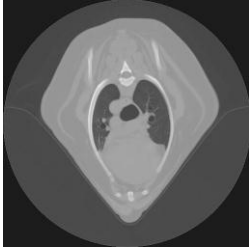
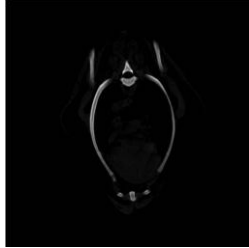
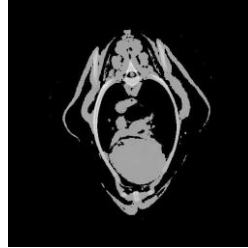
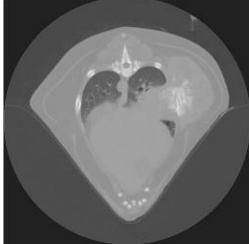
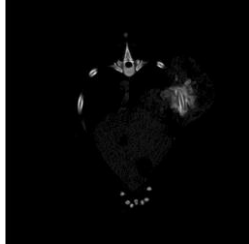
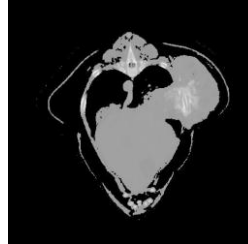
Fig.1. Segmentation process of JPG image

4 Results

The analysis of computation complexity and comparison of implementation speed for the dog rib tumor segmentation shows that our algorithm takes about 1 second. From the visual effect of view by naked eye, the segmentation effect using of CT value is good to using the gray value by the same segmentation algorithm and the same

program implementation. Three examples of segmentation results are shown in Table 2.

Table 2. Three examples of segmentation results

DICOM files	format	Original CT images	Segment by CT value	Segment by gray value
DICOM7_000376				
DICOM7_000378				
DICOM7_000684				

5 Conclusions

We have presented a segmentation approach for CT images, which was demonstrated and validated using two dataset of the same CT image scans, consisting of 312 image volumes. The core of the segmentation is k-means clustering algorithm. The segmentation process consisted of two phases: (1) convert CT value to JPG gray value or not use the original CT value as the data sets for clustering; (2) segmentation bone tissue using the new k-means clustering algorithm program which is implemented with MATLAB 2012a programming language and for two-dimensional data matrix directly.

In this paper, we comprehensively use two values that represent the same object (body tissue), to segment by the same algorithm implementation. The experiment produced strikingly different results. These results may be indicating that the segmentation algorithm of CT image is important and the data for segmentation is

important too. In other words, what kind of data can really represent the real thing? This could be a complex problem.

Acknowledgment

Funds for this research were provided by the Doctor Foundation of Tianjin Agriculture University, No.2012D09.

References

1. KamdarMR,Gomez RA,Ascherman JA: Intracranial volumes in a large series of healthy children[J]. *Plast Reconstr Surg* 124:2072–2075, 2009
2. Mardini S, See LC, Lo LJ, Salgado CJ, Chen YR: Intracranial space,brain, and cerebrospinal fluid volume measurements obtained with the aid of three-dimensional computerized tomography in patients with and without Crouzon syndrome[J]. *J Neurosurg* 103:238–246, 2005
3. Strik HM, et al: Three-dimensional reconstruction and volumetry of intracranial haemorrhage and its mass effect[J]. *Neuroradiology* 47:417–424, 2005
4. Jafarian N, et al.: Automatic fontanel extraction from newborns' CT-images using a model based level set method. Proc. 17th Iranian Conference of Biomedical Engineering (ICBME): Isfahan,3–4 Nov 2010.
5. Kazemi K, et al.: Automatic fontanel extraction from newborns' CT images using variational level set. In: Proceeding CAIP '09 Proceedings of the 13th International Conference on Comp Computer Analysis of Images and Patterns, Berlin, Heidelberg, 2009.
6. A.Meena, K.Raja. K Means Segmentation of Alzheimers Disease in PET scan datasets: An implementation. International Joint Conference on Advances in Signal Processing and Information Technology, SPIT2012, ISSN: 1867-8211 pp. 158-162, 2012.
7. Piyush Valvi, Brijesh Shah, Satish Shah. Improved fuzzy c-mean algorithm for Medical Image segmentation[J]. *International Journal of Engineering Research & Technology (IJERT)*, ISSN: 2278-0181 Vol. 1 Issue 3, May – 2012.
8. Ajala Funmilola A, Oke O.A, Adedeji T.O, Alade O.M, Adewusi E.A. Fuzzy k-c-means Clustering Algorithm for Medical Image Segmentation[J]. *Journal of Information Engineering and Applications*, ISSN 2225-0506 (online), Vol 2, No.6, 2012.
9. Tianhu Lei, Wilfred Sewchand. Statistical approach to X-Ray CT imaging and its applications in image analysis – part II: A new stochastic model-based image segmentation technique for X-Ray CT image [J]. *IEEE Transactions on Medical Image*, 1992, 11(1):62–69.
10. Z. Liang, J.R. MacFall, and D.P. Harrington. Parameter estimation and tissue segmentation from multispectral MR images[J]. *IEEE Transactions on Medical Image*, 1994, 13:441–449.