Region Based Coding of Volume Data Using Finite Mixture Model

Naoyuki Ichimura
Electrotechnical Laboratory
1-1-4 Umezono, Tsukuba,Ibaraki, 305 JAPAN
ichimura@etl.go.jp

Abstract

The region-based coding of the volume data is proposed in the present paper. Since the volume data is the three dimensional one, the visualization method is needed to display it. The region segmentation is indispensable to visualize the desired portions only. In the proposed method, the region segmentation result is utilized to code the volume data; the contours and the colors of the regions are coded. The finite mixture model is used to segment the data, and the parameters of this model is used to represent the opacity needed in the volume rendering. The preliminary results of the experiments for the CT data are shown.

1 Introduction

A volume data is the sampled function of the three spatial dimensions. The typical example is the data generated from a set of CT or MRI images. As we can see from this example, the volume data is useful to represent the internal structure of an object. However, the size of the volume data is normally very large; thus, the coding method for the volume data is needed.

The coding method for the volume data has been developed based on the signal theory such as vector quantization[1], DCT[2] and wavelet[3],[4]. These methods are called waveform-based coding, because the image is treated as a signal waveform. Recently, the coding method which utilizes the structural information extracted from the image has been paid much attention[5][6]. The image is treated as the union of some segments in this method.

The region-based coding is one of the methods using the structural information. In the region-based coding, the image is segmented into the homogeneous regions, and then the contours and the colors of the regions are coded separately. This method has been applied to the gray-scale image, the color image and the image sequence coding [5]-[8].

In the present paper, the region-based coding of the volume data is proposed. There are two reasons in applying the region-based coding to the volume data. One reason is the necessity of region segmentation in the visualization of the volume data. Since the volume data is the three dimensional one, the rendering method which converts the volume data into the two dimensional image is needed. The region segmentation is indispensable to render the desired portions only. That is,

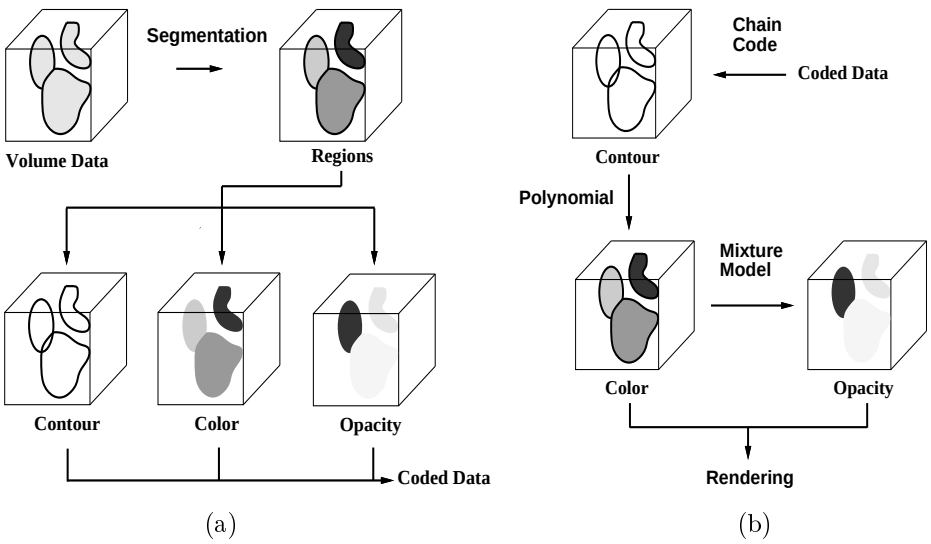

Figure 1: Outline of the proposed coding method. (a)encoding. (b)decoding.

the structural information is inherently needed in this visualization. This information can be utilized to code the data. Another one is the easiness in handling of the volume data. The segments in the volume data can be extracted directly from the coded data. Thus the only part of the volume data can be transmitted and processed easily.

Basically, the proposed method is the extension of the region-based coding for the two dimensional image[5] into the three dimensional data, and the volume rendering[9] is assumed to be used in the visualization of the volume data. Figure 1 shows the outline of the proposed coding method. First, the feature of the each voxel is extracted, then the probability density function of the feature vectors is estimated using the finite mixture model. The volume data is segmented using a posteriori probability of the feature vectors for each component density function of the model.

The finite mixture model has been used in the segmentation of the two dimensional image[10][11]. For the volume data, the use of this model for segmentation has also been presented[12]. However the formula of the component density and the algorithm for parameter estimation of the model were not given. In the present paper, the multivariate t distribution is used as the component and the concrete algorithm for parameter estimation is presented.

The importance of the finite mixture model in the coding of the volume data is that the estimated parameters of the model can be used as the compressed representation of the opacity needed in rendering; because a posteriori probability for each feature vector can be used as the opacity of each voxel.

After the region segmentation, the information of the regions, i.e. the contours and the colors, are coded (Fig.1(a)). The contours of the regions are represented by the eight directional chain codes in each slice. The colors of the regions are

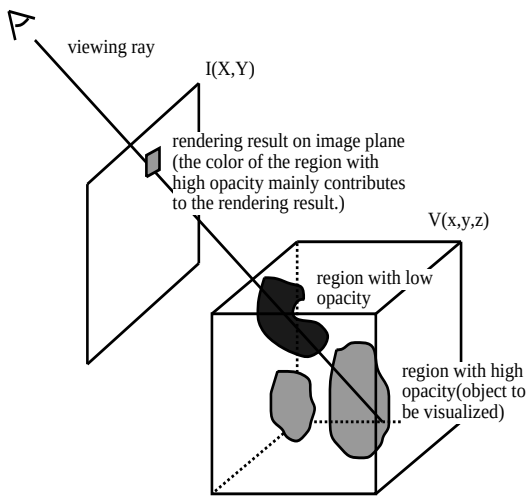

Figure 2: Volume rendering.

approximated by the third order polynomial. Hence the colors can be represented as the coefficients of this polynomial.

The proposed method is applied for the CT data to consider its performance.

2 Volume Rendering

The volume rendering[9] is used as the visualization method in the proposed method. In the volume rendering, the portions to be visualized in the volume data should be segmented to be assigned the high opacity. After the segmentation, the colors of the voxels calculated by the shading process are projected onto the image plane $I(X,Y)$ through the transparencies added from back to front along the viewing ray (Fig.2). This process can be represented as follows:

$$I\left(X,Y\right)=\sum_{i=1}^{K}\left[c_i o_i \prod_{j=i+1}^{K}\left(1-o_j\right)\right] \tag{1}$$

where c_i and o_i are the color and the opacity of the i-th voxel ($i=1,\ldots,K$) which intersects the viewing ray. Obviously, the region segmentation is indispensable to visualize the desired portions in the volume data by the volume rendering.

3 Region Segmentation Using Finite Mixture Model

The region segmentation for the volume rendering is carried out using the finite mixture model. The distribution of a set of the feature vectors $V=\{\boldsymbol{x}_i\}_{i=1}^{N}$

obtained from the volume data is represented by the following finite mixture model.

$$f\left(\boldsymbol{x}\left|\boldsymbol{\Theta}\right.\right)=\sum_{j=1}^{r}\lambda_j\alpha_j\left(\boldsymbol{x}\left|\boldsymbol{\theta}_j\right.\right) \tag{2}$$

where $\boldsymbol{x}$ is a feature vector in $\boldsymbol{R}^n$, r is the number of component densities $\alpha_j\left(\right)$, $\boldsymbol{\Theta}=\left\{\{\lambda_j\}_{j=1}^r\ \ \{\boldsymbol{\theta}_j\}_{j=1}^r\right\}$ is a set of parameters, $\{\lambda_j\}_{j=1}^r$ are the mixing proportions such that $\sum_{j=1}^r\lambda_j=1, 0\leq\lambda_j\leq 1$, and $\boldsymbol{\theta}_j$ is a set of parameters of $\alpha_j\left(\right)$.

The multivariate t distribution is used as a component.

$$\begin{aligned}
\alpha\left(\boldsymbol{x}\left|\boldsymbol{\theta}\right.\right) &= \left|\boldsymbol{V}^{-1}\right| h\left(d_m^2\left|\nu\right.\right) && (3)\\
h\left(d_m^2\left|\nu\right.\right) &= \frac{\Gamma\left\{\left(\nu+n\right)/2\right\}}{\left\{\Gamma\left(1/2\right)\right\}^n\Gamma\left(\nu/2\right)\nu^{n/2}}\left(1+d_m^2/\nu\right)^{-(\nu+n)/2} && (4)\\
d_m^2 &= \left(\boldsymbol{x}-\boldsymbol{m}\right)^t\boldsymbol{V}^{-1}\left(\boldsymbol{x}-\boldsymbol{m}\right) && (5)
\end{aligned}$$

where $\boldsymbol{\theta}=\{\boldsymbol{m},\boldsymbol{V},\nu\}$, and $\boldsymbol{m}$ $(n\times 1)$, $\boldsymbol{V}$ $(n\times n)$ and ν are the location, the scatter matrix and the degree of freedom, respectively. This distribution is identical to the normal distribution when $\nu\rightarrow\infty$, and can represent more heavy tailed distribution by changing ν.

The parameters of the finite mixture model are estimated by a maximum likelihood method. The evaluation function for parameter estimation is as follows:

$$E=\sum_{i=1}^{N}\log\sum_{j=1}^{r}\lambda_j\alpha_j\left(\boldsymbol{x}_i\left|\theta_j\right.\right)-\xi\left(\sum_{j=1}^{r}\lambda_j-1\right) \tag{6}$$

where ξ is a Lagrange multiplier. The EM(Expectation-Maximization) algorithm [13],[14] can be used to maximize above evaluation function. The EM algorithm for the current model is as follows[15]:

<u>[Step 1(E-step)]</u> Calculate a posteriori probability $P\left(S_j\left|\boldsymbol{x}_i\right.\right)$ for the j-th component S_j and the weight $w\left(d_{mi}^2\left|\nu^c\right.\right)$ under the current approximation of the parameters $\boldsymbol{\Theta}^c=\left\{\ \left\{\lambda_j^c\right\}_{j=1}^r,\ \left\{\boldsymbol{\theta}_j^c\right\}_{j=1}^r\right\}$.

$$\begin{aligned}
P\left(S_j\left|\boldsymbol{x}_i\right.\right) &= \frac{\lambda_j^c\alpha_j\left(\boldsymbol{x}_i\left|\boldsymbol{\theta}_j^c\right.\right)}{f\left(\boldsymbol{x}_i\left|\boldsymbol{\Theta}^c\right.\right)} && (7)\\
w\left(d_{mi}^2\left|\nu_j^c\right.\right) &= -2\frac{\partial\log h\left(d_{mi}^2\left|\nu_j^c\right.\right)}{\partial d_{mi}^2}=\frac{\nu+n}{\nu+d_{mi}^2} && (8)\\
d_{mi}^2 &= \left(\boldsymbol{x}_i-\boldsymbol{m}_j^c\right)^t\boldsymbol{V}_j^{c-1}\left(\boldsymbol{x}_i-\boldsymbol{m}_j^c\right) && (9)\\
i &= 1,\ldots,N; j=1,\ldots,r
\end{aligned}$$

<u>[Step 2(M-step)]</u> Calculate the parameters $\boldsymbol{\Theta}^+$.

$$\lambda_j^+ = \frac{1}{N}\sum_{i=1}^{N}P\left(S_j\left|\boldsymbol{x}_i\right.\right) \tag{10}$$

$$\boldsymbol{m}_j^+ = \frac{\sum_{i=1}^N P\left(S_j \left|\boldsymbol{x}_i\right.\right) w\left(d_{mi}^2 \left|\nu_j\right.\right) \boldsymbol{x}_i}{\sum_{i=1}^N P\left(S_j \left|\boldsymbol{x}_i\right.\right) w\left(d_{mi}^2 \left|\nu_j\right.\right)} \tag{11}$$

$$\boldsymbol{V}_j^+ = \frac{\sum_{i=1}^N P\left(S_j \left|\boldsymbol{x}_i\right.\right) w\left(d_{mi}^2 \left|\nu_j\right.\right) \left(\boldsymbol{x}_i - \boldsymbol{m}_j^+\right) \left(\boldsymbol{x}_i - \boldsymbol{m}_j^+\right)^t}{\sum_{i=1}^N P\left(S_j \left|\boldsymbol{x}_i\right.\right)} \tag{12}$$

$$j = 1, \ldots, r$$

[Step 3(M-step)] Calculate the parameter $\left\{\nu_j^+\right\}_{j=1}^r$ using the quasi-Newton method under the fixed ${\lambda_j}^+$, ${\boldsymbol{m}_j}^+$,${\boldsymbol{V}_j}^+$.
[Step 4] If

$$\left|l\left(\boldsymbol{\Theta}^c\right) - l\left(\boldsymbol{\Theta}^+\right)\right| < \epsilon \tag{13}$$

is satisfied or the number of iteration exceeds a given number, then terminate the algorithm. Otherwise, update Θ^c by Θ^+ and go to Step 1.

The maximum likelihood solution for ν can not be represented in the analytical form, so the numerical search in Step 3 is needed.

The weight of eq.(8) shows that the data far from the location of the t distribution have small effect for parameter estimation[15]. That is, the segmentation using this finite mixture model can reduce the effect of the outliers in parameter estimation like as the robust clustering[16].

A posteriori probability given by eq.(7) is used as the opacity of each voxel. The components corresponding to the portions to be coded are selected, and the opacities of all voxels are calculated using them.

The region segmentation is carried out by extracting the voxels which have larger opacity than a given threshold.

4 Coding of Contour, Color and Opacity

After the region segmentation, the information of the regions are coded as follows.

The contours of the regions are represented by the chain code. First, the three dimensional connected components are extracted. And then, the contours of each connected component are represented by the eight directional chain code on each slice. Thus the contours are represented by the slice numbers, the start points of the chains and the chain code numbers.

The colors of each connected components are approximated by the third order polynomial as follows:

$$C = WP \tag{14}$$

$$C = \left[\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N\right] \quad (p \times N), \boldsymbol{c}_i = \left(c_{i1}, \ldots, c_{ip}\right)^t, i = 1, \ldots, N \tag{15}$$

$$W = \left[\boldsymbol{w}_1^t, \ldots, \boldsymbol{w}_p^t\right] \quad (p \times m), \boldsymbol{w}_j = \left(w_{j1}, \ldots, w_{jm}\right)^t, j = 1, \ldots, p \tag{16}$$

$$P = \left[\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N\right] \quad (m \times N) \tag{17}$$

$$\begin{aligned} \boldsymbol{p}_i = {} & (1, x_i, y_i, z_i, x_iy_i, y_iz_i, z_ix_i, x_ix_i, y_iy_i, z_iz_i, x_ix_iy_i, x_ix_iz_i, x_iy_iy_i, \\ & x_iz_iz_i, y_iy_iz_i, y_iz_iz_i, x_iy_iz_i, x_ix_ix_i, y_iy_iy_i, z_iz_iz_i)^t, i = 1, \ldots, N \end{aligned} \tag{18}$$

where N is the number of voxels in the connected component, p is the number of dimensions of the color and (x_i, y_i, z_i) are the coordinates of each voxel. The colors

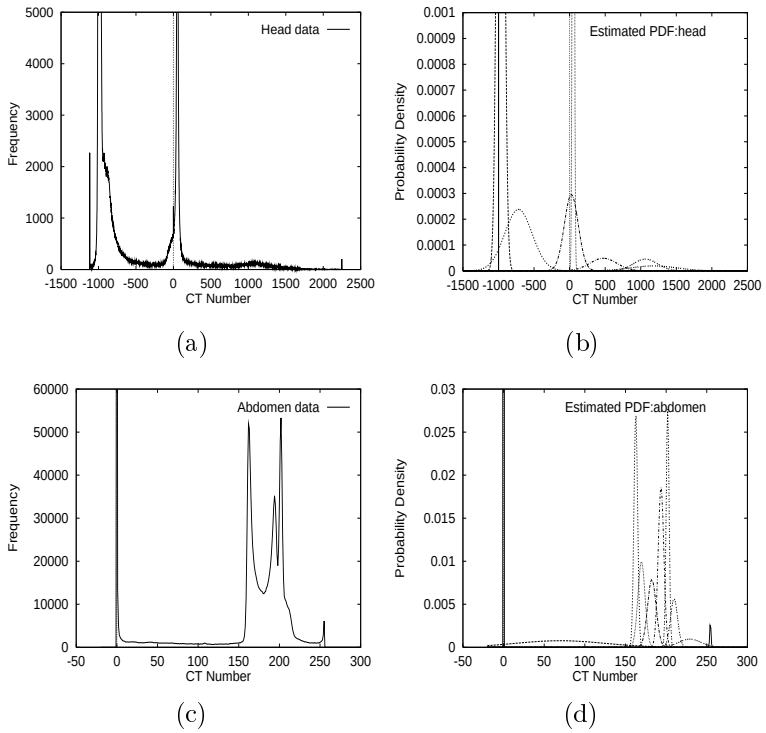

Figure 3: Estimation result of the probability density function for the histogram of the CT number by the finite mixture model. (a) histogram of the CT number:head. (b) estimated probability density function using 8 components. (c) histogram of the CT number:abdomen. (d) estimated probability density function using 10 components.

are represented by the matrix W constructed by m coefficients of the polynomial.

$$W = CP^{+} \tag{19}$$

where P^{+} is the Moore-Penrose generalized inverse matrix of P.

The opacities are represented by the estimated parameters of the finite mixture model. Given these parameters, the opacities can be calculated by eq.(7).

Each information is used in the decoding process as shown in Fig.1(b).

5 Experimental Result

The preliminary results of experiments for the CT data of the head ($128 \times 128 \times 128$,12 [bit/voxel]) and the abdomen ($128 \times 128 \times 94$,8 [bit/voxel]) are shown. The peak-signal-to-noise ratio(PSNR) is used to evaluate the performance of coding. The PSNR is defined by

$$PSNR = 10 \log_{10} \frac{P^2}{\frac{1}{N} \sum_{i=1}^{N} \left(v_i - \hat{v_i}\right)^2} \tag{20}$$

Table 1: Coding performance for the CT data.

Data set	Number of regions	PSNR [dB]	Compression ratio in segmented area	Data size decrement by both coding and segmentation
Head	50	20.9	6.4	99.6
Abdomen	1304	26.9	2.1	18.2

Table 2: Percentage of the information quantity.

Data set	Number of regions	Contour [%]	Color [%]	Opacity [%]
Head	50	97.2	2.3	0.5
Abdomen	1304	90.6	9.2	0.2

where P is the peak signal value, v_i is the intensity of the original volume data and $\hat{v}_i$ is the one of the decoded data.

The feature used in the segmentation was the CT number of each voxel(Fig.3(a) and (c)). The probability density function of the CT number was estimated by the finite mixture model(Fig.3(b) and (d)). The 8 and 10 components were used for the head and the abdomen data, respectively. These components were needed to segment the volume data into the meaningful parts.

In the head data, the components covered the range [-1500,0], [-400,500] and [0,2000] in the histogram mainly corresponded to the air, the soft tissue and the skull, respectively(Fig.3(b)). In this experiment, the skull was selected as the coding object. Hence, the opacities of all voxels were calculated using the components covered the range [0,2000]. Then voxels which had larger opacity than the threshold 0.1 were segmented. Of course, the skull can be segmented by simple thresholding. However, the probability density function of the feature can not be obtained; the opacity can not be calculated from the model.

In the abdomen data, the components covered the range [0,150], [150,200], [190,220] and [200,260] in the histogram mainly corresponded to the air, the skin and muscle, the internal organs and the bone, respectively(Fig.3(d)). The internal organs and the bone were selected as the coding objects. The segmentation was carried out the same manner as the head data.

The skull was correctly segmented by the finite mixture model(Fig.4(a) and (b)). To render the decoded data, the opacities were calculated by eq.(7) using the parameters of the finite mixture model included in the coded data. To improve the quality of the rendering result of the decoded data, the three dimensional Gaussian filter with standard deviation 0.7 was used. The rendering results of the decoded data showed that the major features of the skull were sufficiently reconstructed although the detailed features were lost(Fig.4(c) and (d)). The compression ratio of the skull was 6.4:1 (1.87[bit/voxel])(Table 1). The size of the volume data was reduced to 1/15.5 of the original size by the region segmentation. As a result, the

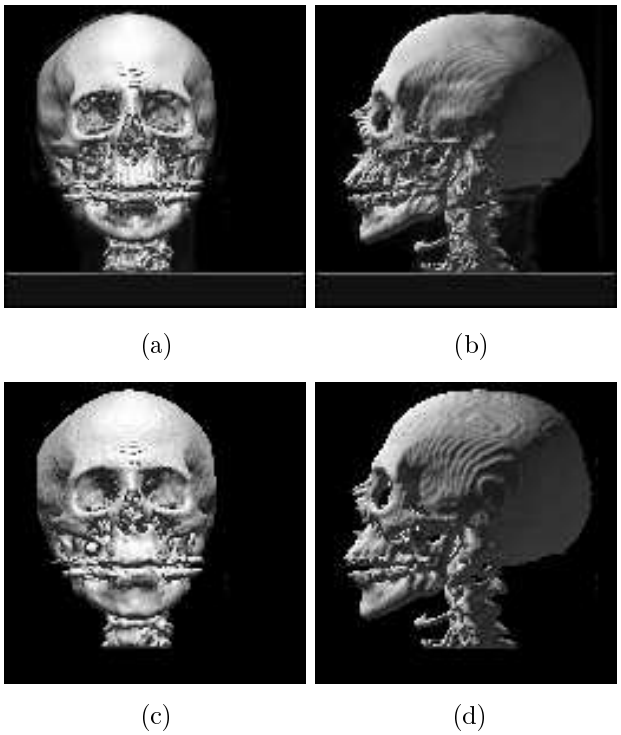

Figure 4: Experimental result for the CT data of the head. (a) original data(front view). (b) original data(side view). (c) decoded data(front view). (d)decoded data(side view).

total reduction of the size of the volume data was 1/99.6(Table 1).

In the abdomen data, there were many small regions in the segmentation result(Fig.5(a) and (b)). The number of regions was 1304. In such case, the efficiency of the region-based coding is decreased. Thus, although the major features of the objects were also sufficiently reconstructed in this case(Fig.5(c) and (d)), the compression ratio was low;2.1:1(3.81[bit/voxel])(Table 1). The decoded image was clearer than the original data, particularly in the bottom view(Fig.5(d)). This is the effect of thresholding of the opacity in the segmentation. The regions which had opacity smaller than 0.1 were not segmented, so semi-transparent regions were not contributed in the rendering result. And the noise reduction by the Gaussian filter also affected the decoded image.

The compression ratio and PSNR of the proposed method were inferior to the ones of the waveform-based method such as DCT[2]. The major reason of the low compression ratio was the lossless coding of the contours. Table 2 shows that over 90% of the information in the coded data was for contours. If more efficient representation, e.g. the triangular patch, is applied to the selected contour points to represent the shape of the regions, the compression ratio may be improved.

The reason of the low PSNR is the poor approximation of the CT number

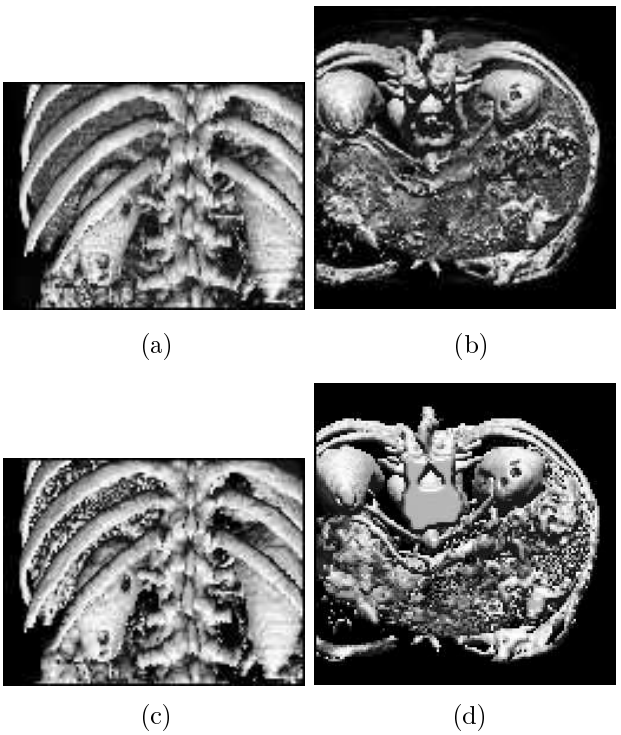

Figure 5: Experimental result for the CT data of the abdomen. (a) original data(front view). (b) original data(bottom view). (c) decoded data.(front view). (d) decoded data(bottom view).

in each region by the polynomial. For this approximation, the waveform-based method may be useful; however, the method which can be applied to the three dimensional data and the arbitrary shaped region is needed.

6 Conclusions

The volume data coding based on the region segmentation has been proposed. The finite mixture model with the t distribution as a component was used to segment the volume data. A posteriori probability could be regarded as the opacity of the voxel. Thus the opacity could be represented as the estimated parameters of the finite mixture model. The contours and the colors of the regions were represented by the chain codes and the coefficients of the polynomial, respectively.

The preliminary results of experiments for the CT data were shown. The major features of the objects were sufficiently reconstructed form the coded data. However, the compression ratio and PSNR were inferior to the waveform-based methods. The following points are the future works for the improvement of the region-based coding performance:(i) the more efficient representation of the region shape; the selection of the significant contour points is important. (ii) the develop-

ment of the waveform-based method which can be applied to the three dimensional data and the arbitrary shaped region.

Acknowledgements

The author would like to thank the members of the Image Understanding Section, Electrotechnical Laboratory for their helpful discussions. Thanks also to the reviewers for their useful comments.

References

[1] P. Ning and L. Hesselink: "Fast Volume Rendering of Compressed Data," Proc. of IEEE Visualization '93, pp.11-18, 1993

[2] B. Yeo and B. Liu: "Volume Rendering of DCT-Based Compressed 3D Scalar Data," IEEE Trans. on Visualization and Computer Graphics, **Vol.1**, No.1, pp.29-43, 1995

[3] S. Muraki: "Volume data and wavelet transforms," IEEE Computer Graphics & Applications, **Vol.13**, No.4, pp.50-56, 1993

[4] S. Muraki: "Multiscale Volume Representation by a DOG Wavelet," IEEE Trans. on Visualization and Computer Graphics, **Vol.1**, No.2, pp.109-116, 1995

[5] M.Kunt,A.Ikonomopoulos and M.Kocher: "Second-Generation Image-Coding Techniques," Proc. IEEE, **Vol.73**, No.4, pp.549-574, 1985

[6] J.Y.A. Wang and E.H. Adelson: "Representing Moving Images with Layers," IEEE Trans. on Image Processing, **Vol.3**, No.5, pp.625-638, 1994

[7] Y. Horita, M. Miyahara, K. Ohtake: "Region Segmentation Coding Using the Local Feature of Color Information," Trans. on the IEICE D-II, **Vol.J76-D-II**, No.5, pp.1023-1037, 1993 (in Japanese)

[8] M. Gilge, T. Engelhardt and R. Mehlan: "Coding of Arbitrarily Shaped Image Segments Based on a Generalized Orthogonal Transform," Signal Processing:Image Communication, **Vol.1**, No.2, pp.153-180, 1989

[9] M. Levoy: "Display of Surfaces from Volume Data," IEEE CG & A, **Vol.8**, No.5, pp.29-37, 1988

[10] S.L. Sclove: "Application of the Conditional Population-Mixture Model to Image Segmentation," IEEE Trans. on Pattern Analysis and Machine Intelligence, **Vol.PAMI-5**, No.4, pp.428-433, 1993

[11] R. Samadani: "A Finite Mixtures Algorithm for Finding Proportions in SAR Images," IEEE Trans. on Image Processing, **Vol.4**, No.8, pp.1182-1186, 1995

[12] R.A. Drebin, L. Carpenter, P. Hanrahan: "Volume Rendering," Computer Graphics, **Vol.22**, No.4, pp.51-58, 1988

[13] A.P. Dempster, N.M. Laird and D.B. Rubin: "Maximum Likelihood from Incomplete Data via the EM Algorithm," J.R.Statist.Soc. B, **Vol.39**, No.1, pp.1-38, 1977

[14] R.A. Redner and H.F. Walker: "Mixture Densities,Maximum Likelihood and the EM Algorithm," SIAM Revies, **Vol.26**, No.2, pp.195-239, 1984

[15] K.L. Lange,R.J.A. Little and J.M.G. Taylor: "Robust Statistical Modeling Using the t Distribution," Journal of the American Statistical Association, **Vol.84**, No.408, pp.881-896, 1989

[16] N. Ichimura: "Inexhaustive Region Segmentation by Robust Clustering," Proc. of IEEE International Conference on Image Processing, **Vol.III**, pp.77-80, 1995