3D Region Merging for Segmentation of Teeth on Cone-Beam Computed Tomography Images

Rarasmaya Indraswari
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
rarasmaya16@mhs.if.its.ac.id

Takio Kurita
Department of Information Engineering
Hiroshima University
Hiroshima, Japan
tkurita@hiroshima-u.ac.jp

Agus Zainal Arifin
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
agusza@cs.its.ac.id

Nanik Suciati
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia

Eha Renwi Astuti
Faculty of Dental Medicine
Universitas Airlangga
Surabaya, Indonesia

Dini Adni Navastara
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia

***Abstract*—Segmentation of teeth in Cone-Beam Computed Tomography (CBCT) images is challenging problem due to its noise and the similar grayscale intensity of bone and teeth element. In this paper we proposed a new method based on three-dimensional (3D) region merging and histogram thresholding for automatic segmentation of teeth on CBCT images. The proposed 3D region merging algorithm can recognized the teeth element that have similar intensity with the bone element based on the three-dimensional (3D) information of the neighboring slices of the CBCT image. Merging the teeth region will lead to more homogenous grayscale intensity distribution inside the teeth. Then histogram thresholding that utilized the characteristic of CBCT images is performed to binarize the grayscale images and obtain the teeth object. The average accuracy, sensitivity, and specificity of the proposed method are 97.75%, 80.22%, and 98.31%, respectively. The proposed method is fully automatic, therefore lead to more objective and reproducible results.**

Keywords—cone-beam computed tomography, region merging, teeth segmentation, three-dimensional image

I. INTRODUCTION

Cone-Beam Computed Tomography (CBCT) scan has been used for wider applications in dentistry due to some limitations of conventional Computed Tomography (CT) scan [1] [2] [3]. CBCT provides volumetric or three-dimensional (3D) reconstruction of the jaw and face area. It gives CBCT advantages of other two-dimensional (2D) imaging techniques of jaw and face area, including panoramic radiographs which is commonly used in dental imaging. Many examination that usually done using the aid of dental panoramic radiographs has been redirected to CBCT because the 3D reconstruction CBCT provides more details to help the examiner do the required analysis.

Segmentation on CBCT usually done to obtained the teeth or jaw bone as object for assisting orthodontal suregery. Segmentation of teeth is also needed is also needed in other applications such as human identification, diagnosing dental syndromes, treatment simulations, etc. [4]. Due to the limitation of the radiation dose, CBCT images may be noisier than Multi-slice Computed Tomography (MSCT) images. And considering that the grayscale intensity of bone and teeth element in CBCT is almost similar, separating the bone from the tooth element become a challenging problem [5] [6].

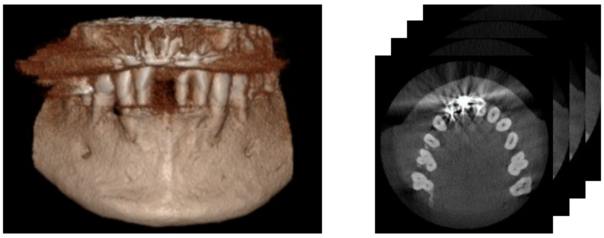
Several image segmentation methods have been proposed to segment teeth object in CBCT images. Those method can be classified into thresholding method, edge-based segmentation method, region-based method, and hybrid method [5] [4]. However, due to the heterogeneous distribution of grayscale intensity inside the teeth, especially if there is appearance of metal artefact, a global thresholding algorithm may leads to under segmentation or over segmentation problem [5] [4]. Active contour methods that based on edge or region need a predetermined ROI (region of interest) or seed points that usually are chosen manually, which is exhaustive to be applied to the 3D images such as CBCT, especially if the method using 2D approach. Moreover, those methods may fail if the boundary between the object and the background is not clear [5]. Hybrid segmentation methods combine the edge information and intensity distribution with the aids of shape priors to overcome those problems. But building shape priors for 3D images that have large variations such as teeth in CBCT images require large training sets [5]. Providing training set, ROI, seed points, etc. is exhaustive for 3D medical image segmentation and it will increase the risk of human error.

In this paper we proposed a method based on 3D region merging and histogram thresholding for fully automatic segmentation of teeth on CBCT images. By utilizing the 3D information from the neighboring slices of the CBCT images, the region merging algorithm can recognized the teeth element that have similar intensity with the bone element and separated them. The histogram thresholding that utilized the characteristic of CBCT images is performed to binarize the grayscale images and obtain the teeth object. By merged the teeth region before the histogram thresholding is performed, the distribution of grayscale intensity inside the teeth will be more homogenous, hence the over segmentation or under segmentation problem that happened can be reduced.

II. PROPOSED METHOD

A. Data

The dataset that has been used in this experiment is CBCT (Cone-Beam Computed Tomography) scan of human's jaw, as shown in Fig. 1. We acquired this scan from the hospital Rumah Sakit Gigi dan Mulut, Universitas Airlangga (RSGM UNAIR). The dataset consisted of jaw images from 6 patients. The segmentation object was teeth and the manually annotated ground truth was confirmed by experts. Each data had sizes of 266 x 266 x 200 voxels. The 2D images were obtained by slicing the 3D image along the axial plane, hence resulting in 200 axial slices that have size of 266 x 266 pixels.



(a) (b)

Fig. 1. CBCT images; (a) Visualization using volumetric rendering; (b) 2D axial slices

B. 3D Region Merging

The 3D region merging method that was used in this research is developed from the Statistical Region Merging (SRM) method proposed by Nock and Nielsen (2004) [7]. The SRM method calculate the gradients of input 2D image I and then merge the smallest gradients into region if their distance satisfy the merging criterion. The process of the proposed 3D region merging method is as follows.

First, the gradients of input 3D image I along the x, y, and z axis (G_x, G_y, and G_z) are calculated using Sobel convolution mask. The 3D kernel H for Sobel convolution process along the x, y, and z direction are shown in (1)-(3), respectively. The gradients of the input image along x, y, and z direction (G_x, G_y, and G_z) is calculated using (4)-(6), respectively, where * denotes the 3D convolution operation. While 2D convolution operation calculate the gradient of a pixel based on its 8-neighborhood along the x and y axis, the 3D convolution operation calculate the gradient of a voxel based on its 26-neighborhood along the x, y, and z axis.

$$H_x(:,:,1) = \begin{bmatrix} -1 & 0 & 1 \\ -3 & 0 & 3 \\ -1 & 0 & 1 \end{bmatrix} \quad H_x(:,:,2) = \begin{bmatrix} -3 & 0 & 3 \\ -6 & 0 & 6 \\ -3 & 0 & 3 \end{bmatrix} \quad H_x(:,:,3) = \begin{bmatrix} -1 & 0 & 1 \\ -3 & 0 & 3 \\ -1 & 0 & 1 \end{bmatrix} \tag{1}$$

$$H_y(:,:,1) = \begin{bmatrix} -1 & -3 & -1 \\ 0 & 0 & 0 \\ 1 & 3 & 1 \end{bmatrix} \quad H_y(:,:,2) = \begin{bmatrix} -3 & -6 & -3 \\ 0 & 0 & 0 \\ 3 & 6 & 3 \end{bmatrix} \quad H_y(:,:,3) = \begin{bmatrix} -1 & -3 & -1 \\ 0 & 0 & 0 \\ 1 & 3 & 1 \end{bmatrix} \tag{2}$$

$$H_z(:,:,1) = \begin{bmatrix} -1 & -3 & -1 \\ -3 & -6 & -1 \\ -1 & -3 & -1 \end{bmatrix} \quad H_z(:,:,2) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad H_z(:,:,3) = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 1 \end{bmatrix} \tag{3}$$

$$G_x = H_x * I \tag{4}$$

$$G_y = H_y * I \tag{5}$$

$$G_z = H_z * I \tag{6}$$

Those gradients are then merged in a matrix and sorted in increasing order. Small gradient means that two adjacent pixels have high similarity. Let each voxels in the input image I become region and (a, b) be the pair of voxels that has the smallest gradient $G(a, b)$ in the gradient matrix. The distance between the region that contains voxel a (C_a) and the region that contains voxel b (C_b) is calculated using (7)

$$d(C_a, C_b) = |\mu_{C_a} - \mu_{C_b}|, \tag{7}$$

where $d(C_a, C_b)$ is the distance between region C_a and region C_b, μ_{C_a} is the mean intensity of region C_a, and μ_{C_b} is the mean intensity of the region C_b.

If the distance $d(C_a, C_b)$ satisfy the merging criterion, region C_a and C_b are merged into one region ($C_{a\cup b}$), as in (8)

$$C_{a\cup b} = \begin{cases} true \ ; & d(C_a, C_b) \le \sqrt{s^2(C_a) + s^2(C_b)} \\ false \ ; & otherwise \end{cases}, \tag{8}$$

where $s(C_a)$ and $s(C_b)$ is the merging threshold for each region. The merging threshold for region C ($s(C)$) is calculated using (9)

$$s(C) = g\sqrt{\frac{1}{2Qn_C}\ln\left(\frac{\mathrm{R}_{n_C}}{\delta}\right)}, \tag{9}$$

where g is the number of gray level in the image ($g = 256$), Q is the coarseness level ($Q = \{1, 2, 4, \ldots, 256\}$), n_C is the number of voxels in region C, R_{n_C} is the set of regions with n_C voxels, $\delta = \frac{1}{6{n_I}^2}$, and n_I is the number of voxels in the input image I. After that, mean intensity of the merged region are calculated and the smallest gradient $G(a, b)$ is eliminated from the gradient matrix. Those processes are repeated until the gradient matrix is empty.

C. Teeth Segmentation

The steps of the proposed method is shown in Fig. 2. Using the 2D axial slices $I_{axial}(x, y)$ of the CBCT images as the input, a 3D matrix *I(x, y, z)* is formed by stacking all the slices along the z-axis sequentially. Then 3D region merging with the predetermined coarseness level (Q) is performed on the 3D matrix. In this paper the Q value is set to 256 to obtain the coarsest region merging result because there is high similarity of the grayscale intensity between teeth and bone element may. Smaller Q value means a smoother region merging result; hence the teeth and bone element that have similar grayscale intensity may be merged. The result of 3D region merging is a grayscale 3D matrix which intensities have been merged into several clusters.

The next step is determine which clusters are belong to the teeth element using histogram thresholding. There are four main elements in CBCT images, which are air, soft tissue, bone, and teeth element [8]. The air element have the lowest grayscale intensity, followed by soft tissue, bone, and teeth (crown) element, respectively. If the grayscale intensity of the input axial slice is modeled into a histogram, as shown in Fig. 3, it can be seen that there are several modes in the histogram (multimodal). Each mode represent one or more of the CBCT elements. The grayscale histogram of CBCT images can be modeled as the mixture of up to five Gaussian distribution that represents its elements [9]. Hence, in this research, the

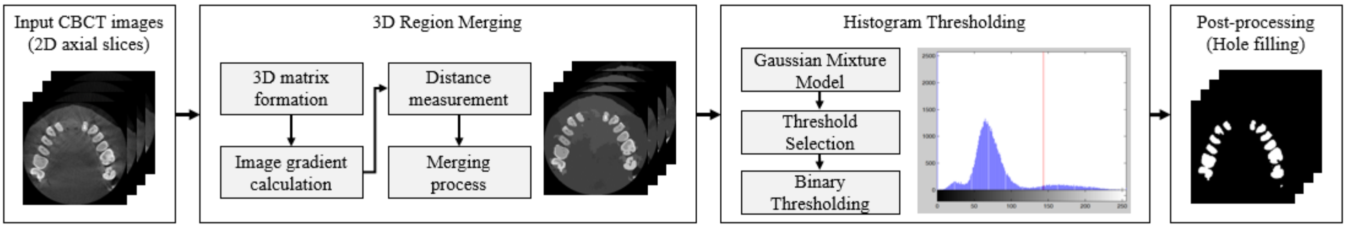


Fig. 2. Proposed method

elements of CBCT images were separated using histogram thresholding based on Gaussian Mixture Model (GMM).

The Expectation-Maximization (EM) algorithm is utilized to predict the GMM of the histogram. Let k be the number of Gaussian distributions in the data. For each data point, the responsibility of each Gaussian distribution is calculated using the predicted mean and standard deviation of each Gaussian distribution. Then using the responsibility value, the mean and standard deviation of each Gaussian distribution will be updated. This process is repeated until it converges.

In this research, the histogram of the input axial slices is separated into four clusters ($k = 4$) according to the number of elements in CBCT images. The grayscale intensity value that become the boundary of the third and fourth cluster is then selected as the threshold to separate the bone and teeth element in the image. After the threshold value of all the input 2D axial slices have been obtained, the final threshold value for the 3D matrix is calculated by averaging the threshold value of all the axial slices. To obtain the binary image of teeth segmentation result, all the grayscale value in the 3D region merging result that below the final threshold is converted to black while all the grayscale value in the 3D region merging result that over the final threshold is converted to white.

Due to the dark color of the root, the histogram thresholding that was performed will not include the teeth root as the segmentation object. Therefore a morphological operation, hole filling, is performed to obtain the teeth root as the segmentation object. Hole filling algorithm check whether there are 'holes' inside the objects in binary image and then 'fill' it, hence the 'holes' also becoming part of the object.

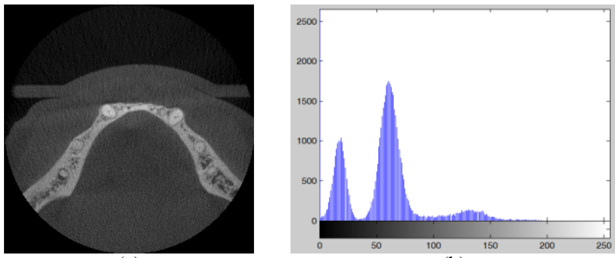
III. Result and Discussion

A. Performance Measurement

The performance of the proposed method is measured using accuracy, sensitivity (true positive rate), and specificity (true negative rate) evaluation metrics. Sensitivity measures the percentage of the correctly classified positive pixels or object. And specificity measures the percentage of the correctly classified negative pixels or background. Sensitivity is not measured on the slices that contains no object. The average accuracy, sensitivity, and specificity of the proposed method are 97.75%, 80.22%, and 98.31%, respectively, as shown in Table I.

B. Effect of Coarseness Level

The coarseness level (Q) is used in the merging criterion that determines when the merging process will be stopped, hence it helps determine how many regions will be left on the region merging result. The smaller the Q value, the smoother the region merging result will become, as in Fig. 4, that shows the region merging result using different Q values on image



(a) (b)

Fig. 3. Axial slice of CBCT image; (a) Grayscale image; (b) Intensity histogram

TABLE I. Performance Measurement of the Proposed Method

Image No.	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	97.07	76.14	97.87
2	98.15	69.25	98.75
3	98.62	75.13	99.27
4	98.05	91.25	98.24
5	97.07	94.28	97.14
6	97.53	75.26	98.57
Mean	**97.75**	**80.22**	**98.31**

no. 1. In this experiment, the effect of coarseness level on the performance of the proposed method is inspected. Table II shows the performance of the proposed method using different Q values on image no. 1. It can be seen that although there is no clear pattern on the accuracy and specificity values, the smaller the Q value, the sensitivity value tends to decrease.

When $Q = 1$, the 3D region merging algorithm merged the input CBCT image into mainly 2 clusters. The first cluster consist of air and soft tissue element, while the second cluster consist of bone and teeth element. This is shown in Fig. 5(j) where the axial slice is located on the lower part of the jaw (slice no. 95 from 200), hence there is bone and teeth element on the image that are merged by the 3D region merging algorithm. Therefore, if the separation between bone and teeth element is not needed, the proposed 3D region merging algorithm may be used for the binary segmentation process directly.

C. Effect of Histogram Thresholding

The histogram thresholding is used for the binarization of the 3D region merging result. With the wrong threshold, even if the region merging algorithm has been correctly grouped the elements of the image, the binary segmented image will also be wrong. Hence it is important to select the appropriate threshold for the teeth segmentation. In this research, a threshold based on the characteristic of CBCT images is selected to perform teeth segmentation. The intensity histogram of the 2D axial slices of the CBCT image is modeled as the mixture of Gaussian distribution with the number of distribution $k = 4$ according to the number of elements in CBCT image (air, soft tissue, bone, and teeth).

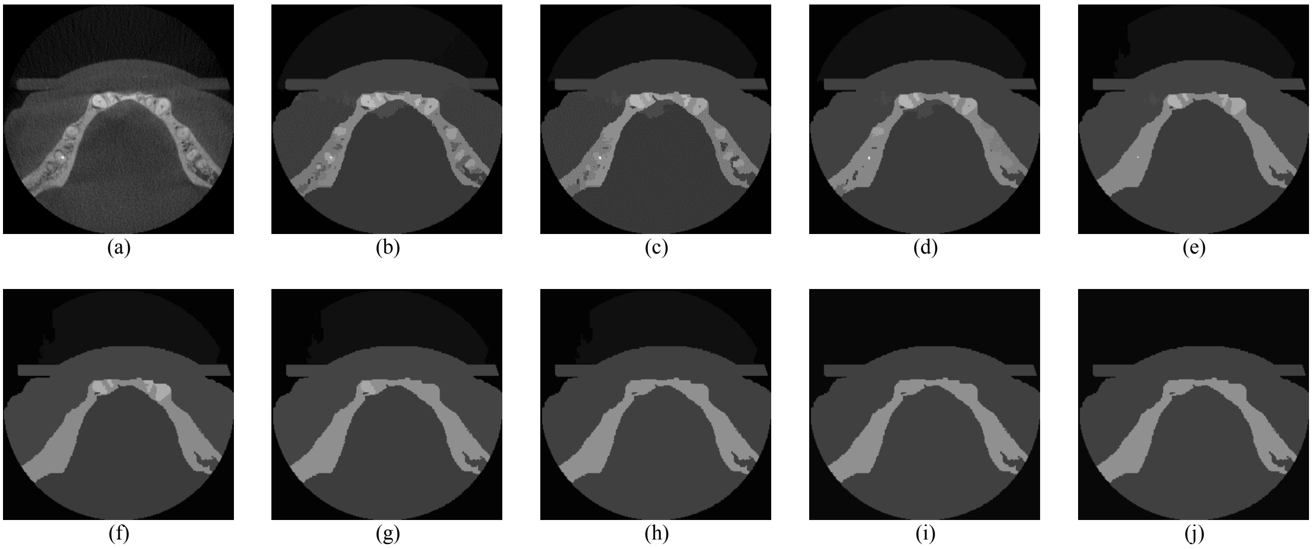


Fig. 4. The effect of coarseness level in the proposed 3D region merging; (a) Input image; (b) $Q = 256$; (c) $Q = 128$; (d) $Q = 64$; (e) $Q = 32$; (f) $Q = 16$; (g) $Q = 8$; (h) $Q = 4$; (i) $Q = 2$; (j) $Q = 1$

TABLE II. THE EFFECT OF COARSENESS LEVEL IN THE PROPOSED METHOD

Q value	Accuracy (%)	Sensitivity (%)	Specificity (%)
256	97.07	76.14	97.87
128	98.47	72.66	99.21
64	97.84	69.02	98.75
32	98.50	64.05	99.49
16	98.29	58.28	99.54
8	97.70	32.98	99.77
4	97.45	21.79	99.81
2	97.50	23.08	99.80
1	97.53	23.65	99.80

The threshold to segment teeth element is the grayscale intensity that become the boundary between bone and teeth element (third border).

However, because not all of the CBCT slices contain those four elements, the appropriate number of distribution k may be different for each slices. The example of this problem is shown in Fig. 5, in which the axial slice did not contain bone element. Hence, the selection of $k = 4$ will be inappropriate and if the thresholding process is done by using the grayscale value of the third border (third red line in Fig. 5(d)) will result in under segmentation problem. To avoid this problem, mean intensity of the selected threshold (third border) of each axial slices is used as the threshold to segment all of the slices in the CBCT image. Therefore, the threshold is not selected individually for each 2D slices, but selected by averaging each individual threshold in a CBCT data. This way, the threshold will still be adaptive for each 3D CBCT data but not too affected by the difference in each slices. However, for future works it is better to develop a method to determine the appropriate number of Gaussian distribution in each slices, so that the performance of the histogram thresholding for teeth segmentation can be increased.

D. Comparison with Statistical Region Merging

The performance measurement of the SRM method for teeth segmentation is shown in Table III, where the average accuracy, sensitivity, and specificity values are 97.79%, 69.00%, and 98.48%, respectively. Compared to the performance of the 3D region merging method that shown in Table I, the accuracy and specificity of the SRM method is a little bit higher than the 3D region merging method, while the sensitivity of the SRM method is lower than the 3D region merging method. This is because the background in CBCT images is more dominant than the object, especially because there is several slices that contains no object (teeth element).

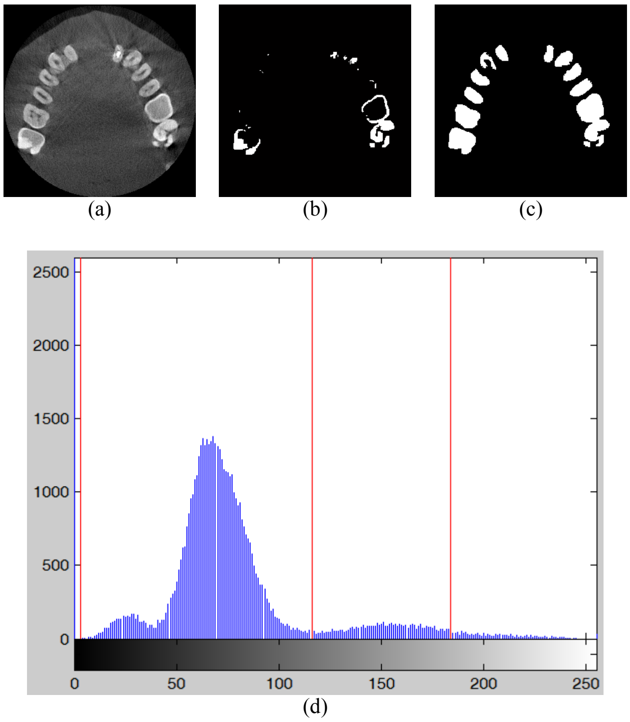


Fig. 5. Histogram thresholding; (a) Input axial slice; (b) Thresholding result based on the threshold value of individual slice; (c) Thresholding result based on average threshold value; (d) Histogram and borders of elements of the input axial slice

Higher sensitivity means that proposed method is better SRM method in detecting the object. By incorporating the information from neighboring slices, the 3D region merging can detect teeth element that have similar intensity with the bone element. For example, in the middle part of the jaw there is no bone element so that the teeth element can be detected rather easily. While in the lower part of the jaw, the teeth element is almost disappeared (only the root remained) and surrounded by the bone element that have similar intensity with the teeth. Therefore, using the 3D region merging, the information of the teeth on the middle part of the jaw can be utilized to help the segmentation process in the lower part of the jaw. The comparison of the proposed 3D region merging method and the SRM method is shown in Fig. 6. It can be seen that the proposed method can still separate the teeth from the bone while the SRM method merged most of the teeth with the bone.

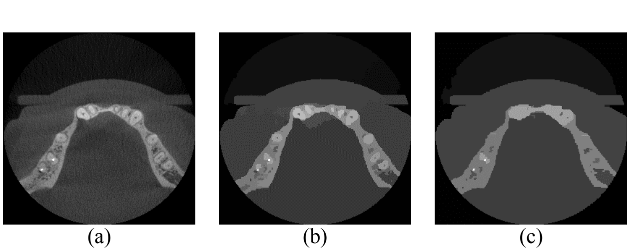
IV. CONCLUSION

CBCT scan has been used for wider applications in dentistry. Segmentation of teeth in CBCT images is challenging problem due to its noise and the similar grayscale intensity of bone and teeth element. In this paper we proposed a method based on 3D region merging and histogram thresholding for automatic segmentation of teeth on CBCT images. The 3D region merging method that was used in this research is developed from the SRM method proposed by Nock and Nielsen (2004). The coarseness level determine when the region merging process will be stopped. By utilizing the 3D information from the neighboring slices of the CBCT images, the region merging algorithm can recognized the teeth element that have similar intensity with the bone element and separated them.

There are four main elements in CBCT images, which are air, soft tissue, bone, and teeth element. These elements is represented in the grayscale intensity histogram of CBCT slices as mixture of Gaussian distributions. Using, EM algorithm to predict the Gaussian Mixture Model of the histogram, threshold that separates the teeth element from the others can be determined. Merging the teeth region before the histogram thresholding is performed will lead the distribution of grayscale intensity inside the teeth to be more homogenous. Therefore, the over segmentation or under segmentation problem that happened can be reduced. To obtain the root of the teeth that have low grayscale intensity, post-processing using hole filling algorithm is performed. The average accuracy, sensitivity, and specificity of the proposed method are 97.75%, 80.22%, and 98.31%, respectively. The proposed method is fully automatic, therefore lead to more objective and reproducible results.

TABLE III. PERFORMANCE MEASUREMENT OF STATISTICAL REGION MERGING

Image No.	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	97.54	61.59	98.40
2	98.03	52.72	98.86
3	98.19	53.56	99.40
4	98.45	90.02	98.67
5	96.58	91.05	96.66
6	97.96	65.04	98.86
Mean	**97.79**	**69.00**	**98.48**



(a) (b) (c)

Fig. 6. Method comparison; (a) Input image; (b) Proposed 3D region merging; (c) Statistical Region Merging

ACKNOWLEDGMENT

This work is supported by the Ministry of Research, Technology, and Higher Education, Indonesia with grant number 128/SP2H/PTNBH/DRPM/2018. We are deeply grateful for PMDSU (*Program Magister Menuju Doktor untuk Sarjana Unggul*) and PKPI (*Program Peningkatan Kualitas Publikasi Ilmiah*) sandwich-like program that enabled this joint-research between Institut Teknologi Sepuluh Nopember (ITS) and Hiroshima University.

REFERENCES

[1] J. B. Ludlow, L. E. Davies-Ludlow, S. L. Brooks and W. B. Howerton, "Dosimetry of 3 CBCT devices for oral and maxillofacial radiology: CB Mercuray, NewTom 3G and i-CAT," *Dentomaxillofacial Radiology,* vol. 35, no. 4, pp. 219-226, 2006.

[2] D. A. Tyndall and S. Rathore, "Cone-beam CT diagnostic applications: caries, periodontal bone assessment, and endodontic applications," *Dental Clinics of North America,* vol. 4, no. 825-841, p. 52, 2008.

[3] B. Al-Rawi, B. Hassan, B. Vandenberge and R. Jacobs, "Accuracy assessment of three-dimensional surface reconstructions of teeth from cone beam computed tomography scans," *Journal of oral rehabilitation,* vol. 37, no. 5, pp. 352-358, 2010.

[4] M. Hosntalab, R. A. Zoroofi, A. A. Tehrani-Fard and G. Shirani, "Segmentation of teeth in CT volumetric dataset by panoramic projection and variational level set," *International Journal of Computer Assisted Radiology and Surgery,* vol. 3, pp. 257-265, 2008.

[5] D. X. Ji, S. H. Ong and K. W. C. Foong, "A level-set based approach for anterior teeth segmentation in cone beam computed tomography images," *Computers in biology and medicine,* vol. 50, pp. 116-128, 2014.

[6] H. Gao and O. Chae, "Individual tooth segmentation from CT images using level set method with shape and intensity prior," *Pattern Recognition,* vol. 43, pp. 2406-2417, 2010.

[7] R. Nock and F. Nielsen, "Statistical region merging," *IEEE Transactions on pattern analysis and machine intelligence,* vol. 26, no. 11, pp. 1452-1458, 2004.

[8] Y. Gan, Z. Xia, J. Xiong, Q. Zhao, Y. Hu and J. Zhang, "Toward accurate tooth segmentation from computed tomography images using a hybrid level set model," *Medical physics,* vol. 42, no. 1, pp. 14-27, 2015.

[9] M. Loubele, F. Maes, F. Schutyser, G. Marchal, R. Jacobs and P. Suetens, "Assessment of bone segmentation quality of cone-beam CT versus multislice spiral CT: a pilot study," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology,* vol. 102, no. 2, pp. 225-234, 2006.