

Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation

Ming-Ni Wu¹

¹Department of Computer
Science and Information
Engineering, National Chung
Cheng University, Chiayi,
Taiwan 621, R. O. C.
mnwu@cs.ccu.edu.tw

Chia-Chen Lin²

²Department of Computer
Science and Information
Management, Providence
University, Taichung, Taiwan
433, R. O. C.
mhl3@pu.edu.tw

Chin-Chen Chang¹³

³Department of Information
Engineering and Computer
Science, Feng Chia University,
Taichung, Taiwan, 40724,
R.O.C
ccc@cs.ccu.edu.tw

Abstract

In this paper, we propose a color-based segmentation method that uses the K-means clustering technique to track tumor objects in magnetic resonance (MR) brain images. The key concept in this color-based segmentation algorithm with K-means is to convert a given gray-level MR image into a color space image and then separate the position of tumor objects from other items of an MR image by using K-means clustering and histogram-clustering. Experiments demonstrate that the method can successfully achieve segmentation for MR brain images to help pathologists distinguish exactly lesion size and region.

1. Introduction

Imaging is an essential aspect of medical science to visualize the anatomical structures of the human body [3, 4]. Several new complex medical imaging modalities, such as X-ray, magnetic resonance imaging (MRI), and ultrasound, strongly depend on computer technology to generate or display digital images. With computer techniques, multidimensional digital images of physiological structures can be processed and manipulated to help visualize hidden diagnostic features that are otherwise difficult or impossible to identify using planar imaging methods.

Segmentation is an important process in most medical image analysis and classification for radiological evaluation or computer-aided [3] diagnosis. Basically, image segmentation methods can be classified into three categories: edge-based methods, region-based methods [6], and pixel-based methods.

K-means clustering is a key technique in pixel-based methods. Because pixel-based methods based on K-means clustering are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable. Furthermore, K-means clustering is suitable for biomedical image segmentation as the number of clusters is usually known for images of particular regions of the human anatomy. Many researchers have proposed related research into K-means clustering segmentation [1, 5]. The improvements achieved by [1, 5] have been remarkable, but more computational complexity and extra software functionality are required. In this paper, we carefully select the appropriate features from brain images as the clustering features to achieve good segmentation results while maintaining the low computation aspect of the segmentation algorithm. Because the color space transformation function in our proposed method is a fundamental operation for most image processing systems, the color space translation does not cause extra overhead in the proposed scheme. Therefore, by using color-based segmentation with K-means clustering to magnetic resonance (MR) brain tumors, the proposed image tracking method maintains efficiency. The experimental results also confirm that the proposed method helps pathologists distinguish exact lesion sizes and regions.

The rest of this paper is organized as follows. Section 2 introduces the basic theorem of histogram statistics segmentation and the K-means clustering method. Our proposed scheme is presented in Section 3. Experimental results are illustrated and discussed in Section 4. Finally, concluding remarks are given in Section 5.

2. Review of histogram statistics and K-means clustering

Two pixel-based segmentation methods are applied in our proposed method. One is histogram statistics and the other is K-means clustering. The histogram method defines single or multiple thresholds to classify an image pixel-by-pixel. A simple approach to determine the gray value threshold T is by analyzing the histogram for peak values and finding the lowest point, which is typically located between two consecutive peak values of the histogram. If a histogram is clearly bi-modal, the histogram statistics method can provide good results. By comparing the gray value of each pixel with the determined threshold T , a pixel can be classified into one or the two classes. An image $f(x,y)$ can be segmented into two classes using a gray value threshold T so that

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \leq T, \end{cases} \quad (1)$$

where $g(x,y)$ is the segmented image with two classes of binary values, “1” and “0”, and T is the threshold assigned to the lowest point, which is located between two peak values of the histogram.

K-means is a widely used clustering algorithm to partition data into k clusters. Clustering is the process for grouping data points with similar feature vectors into a single cluster and for grouping data points with dissimilar feature vectors into different clusters. Let the feature vectors derived from l clustered data be $X = \{x_i | i = 1, 2, \dots, l\}$. The generalized algorithm initiates k cluster centroids $C = \{c_j | j = 1, 2, \dots, k\}$ by randomly selecting k feature vectors from X . Later, the feature vectors are grouped into k clusters using a selected distance measure such as Euclidean distance so that

$$d = \|x_i - c_j\|. \quad (2)$$

The next step is to recompute the cluster centroids based on their group members and then regroup the feature vectors according to the new cluster centroids. The clustering procedure stops only when all cluster centroids tend to converge.

3. Proposed method

In the proposed method, we combine histogram statistics and K-means clustering to track the tumor objects in MR brain images. The flowchart of our method is shown as Fig. 1.

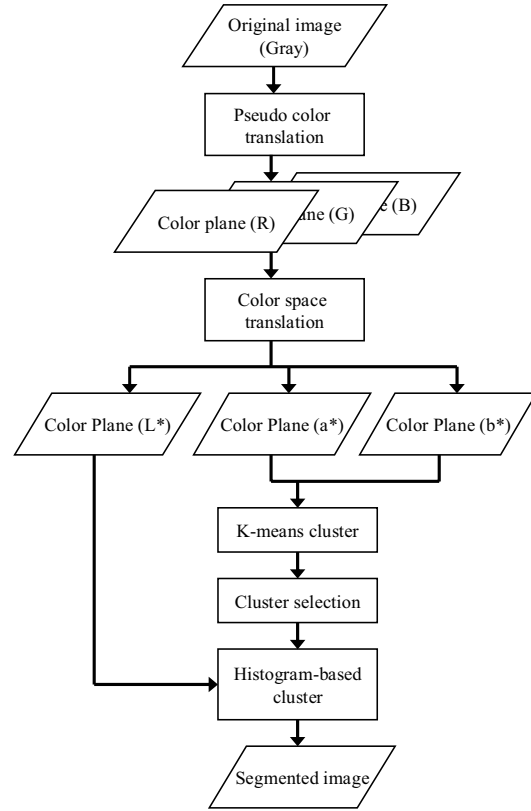


Fig. 1. Block diagram of proposed method

Basically, feature space selection is a key issue in K-means clustering segmentation. The original MR brain image is rendered as a gray-level image that is insufficient to support fine features. To obtain more useful feature and enhance the visual density, the proposed method applies pseudo-color transformation, a mapping function that maps a gray-level pixel to a color-level pixel by a lookup table in a predefined color map. An RGB color map contains R, G, and B values for each item. Each gray value maps to an RGB item. The proposed method has adopted the standard RGB color map, which gradually maps gray-level values 0 to 255 into blue-to-green-to-red color.

To retrieve important features to benefit the clustering process, the RGB color space is further converted to a CIE Lab color model ($L^*a^*b^*$) [2]. The $L^*a^*b^*$ space consists of a luminosity layer L^* , a chromaticity-layer a^* , which indicates where color falls along the red-green axis, and a chromaticity-layer b^* , which indicates where the color falls along the blue-yellow axis. The translating formula calculates the tri-stimulus coefficients first as

$$\begin{aligned}
W &= 0.4303R + 0.3416G + 0.1784B, \\
Y &= 0.2219R + 0.7068G + 0.0713B, \\
Z &= 0.0202R + 0.1296G + 0.9393B.
\end{aligned} \tag{3}$$

The CIELab color model is calculated as

$$\begin{aligned}
L^* &= 116(h(\frac{Y}{Y_s})) - 16, \\
a^* &= 500(h(\frac{W}{W_s}) - h(\frac{Y}{Y_s})), \\
b^* &= 200(h(\frac{Y}{Y_s}) - h(\frac{Z}{Z_s})), \\
h(q) &= \begin{cases} \sqrt[3]{q} & q > 0.008856 \\ 7.787q + \frac{16}{116} & q \leq 0.008856, \end{cases}
\end{aligned} \tag{4}$$

where Y_s , W_s , and Z_s are the standard stimulus coefficients.

Both the a^* and b^* layers contain all required color information. Therefore, the proposed method then classifies the colors in the a^*b^* space using K-means clustering. After the clustering process, the cluster containing an area of interest is selected as the primary segment. To eliminate the pixels which are not related to the interest in the selected cluster, histogram clustering is applied by luminosity feature L^* to derive the final segmented result.

4. Experimental Results

In the proposed method, we convert a gray-level MR brain image into an RGB color image first and then convert the RGB color image into a CIELab color model. Therefore, colors in both the a^* and b^* spaces are feature vectors for K-means clustering. To demonstrate the detection performance of the proposed method, an MR brain image (288×288) containing the pathological change area shown in Fig. 2(a) served as a test image. The RGB color image converted from the gray-level image is shown in Fig. 2(b). To prove that the feature vectors proposed by our method really can provide better segmentation performance, two different data sets were prepared: the gray feature vectors of the original MR brain image and the RGB features derived from the converted RGB color image. Generally, an MR brain image consists of regions that represent the bone, soft tissue, fat and background. In these two different data sources, visual judgments from the gray and color test images suggest three primary clusters in the test image shown in Fig. 2(a) when $k=3$. Figs. 3(a)–(c) show the image labeled by cluster index

from the K-means process for different kinds of feature vectors. Using index labels, we can separate objects in the brain image by three colors: white, gray, and black. The final segmentation results generated by histogram clustering are shown in Figs. 3(d)–(f). By combining Fig. 3(a) and Fig. 3(d), we can see that not only a tumor (in the right half of the image) is identified but also the white matter, cerebrospinal fluid, and the ventricles are. In other words, the segmentation result cannot exactly identify the position of the tumor shown in Fig. 3(d). The same phenomenon exists in Fig. 3(e). However, the segmentation result generated by the proposed method can ignore most of the white matter, cerebrospinal fluid, and ventricles while exactly locating the position of the tumor. Therefore, the segmentation performance of the proposed features derived from the CIELab color model and the proposed method is confirmed.

5. Conclusions

In this paper, a color-based segmentation method based on K-means clustering for tracking tumor in the MRI brain image is proposed. A preliminary experiment conducted on the MRI brain image demonstrates encouraging results. By using the features derived from CIELab color model can provide good segmentation performance with the proposed method and the location of a tumor or lesion can be exactly separated from the colored image. In addition, the proposed method simply combines color translation, K-means clustering and histogram-clustering, thus making it efficient and very easy to implement.

References

- [1] Chen, C. W., Luo, J., Parker, K. J., "Image Segmentation via Adaptive K-mean Clustering and Knowledge-based Morphological Operations with Biomedical Applications," *IEEE Transactions on Image Processing*, Vol. 7, No. 12, Dec. 1998, pp. 1673-1683.
- [2] Christine, C., Thomas, F., "A Study of Efficiency and Accuracy in the Transformation from RGB to CIELAB Color Space," *IEEE Transactions on Image Processing*, Vol. 6, No. 7, July 1997, pp. 1046-1048.
- [3] Dhawan, A. P., "A Review on Biomedical Image Processing and Future Trends," *Computer Methods and Programs in Biomedicine*, Vol. 31, No.3-4, 1990, pp.141-183.
- [4] Gonzalez, R. C.; Woods, R. E., *Digital Image Processing*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 2002.
- [5] Ng, H. P., Ong, H. H., Foong, K. W. C., Goh, P. S., Nowinski, W. L., "Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm,"

IEEE Southwest Symposium on Image Analysis and Interpretation, 2006, pp. 61-65.

[6] Tsai, C. S., Chang, C. C., "An Improvement to Image Segment Based on Human Visual System for Object-based

Coding," *Fundamenta Informaticae*, Vol. 58, No. 2, 2004, pp. 167-178.

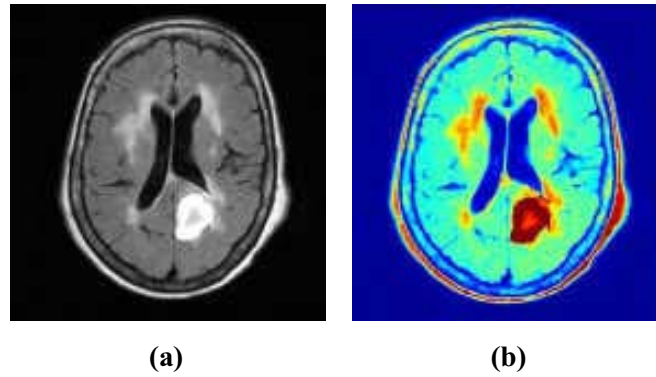


Fig. 2. Original MR brain images. (a) gray-level image, (b) color image.

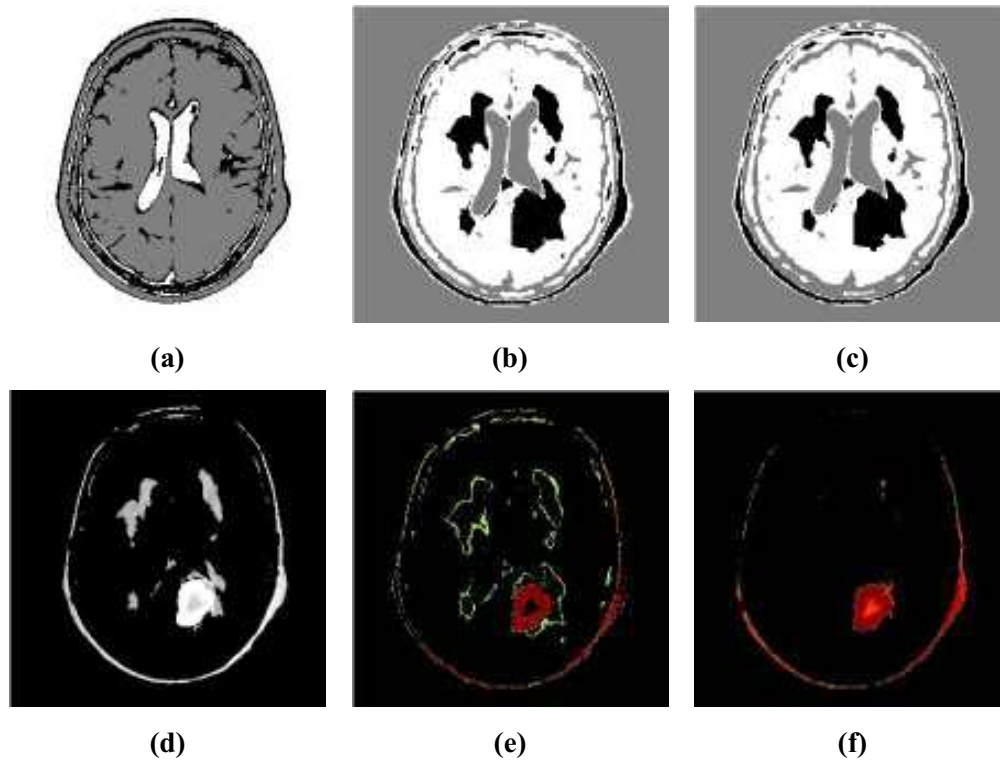


Fig. 3. Comparison of segmentation results. (a) image labeled by cluster index of a gray image; (b) image labeled by cluster index of an RGB image; (c) image labeled by cluster index of the proposed method; (d) final segmentation of (a); (e) final segmentation of (b); (f) final segmentation of (c).