IMPLEMENTATION OF SPATIAL FUZZY CLUSTERING IN DETECTING LIP ON COLOR IMAGES

Agus Zainal Arifin [1], Adhatus Sholichah [2], Anny Yuniarti [3], Dini Adni Navastara [4], Wijayanti Nurul Khotimah [5]
[1,2,3,4,5] Department of Informatics, Faculty of Information Technology
Institut Teknologi Sepuluh Nopember, 60111
agusza@its-sby.edu [1], {licha [2], anny [3]}@cs.its.ac.id, {dini_navastara [4], wijayanti [5]}@if.its.ac.id

ABSTRACT

Dynamic visual information from the lip movement can significantly improve the accuracy and robustness of an automatic speech recognition system in a noisy environment. Useful geometric information about lip movement, such as the temporal variation of mouth width and height, can be obtained easily from a segmented lip. However, there is a difficulty in lip segmentation due to the weak color contrast between lip and face regions. As the result, successful segmentation of lip images cannot be obtained.

Spatial Fuzzy Clustering algorithm (SFCM) adapts conventional FCM. This algorithm is able to take into account both the distributions of data in feature space and the spatial interactions between neighboring pixels during clustering. In SFCM color components are not treated independently. By appropriate preprocessing and postprocessing utilizing the color and shape properties of lip region, successful segmentation of lip images can be reached.

Experiment of SFCM is done by using COLORFERET database. Experiment using different size of neighborhood shows average accuracy 90.9%.

Keywords : segmentation, color lip images, local spatial interaction, spatial fuzzy clustering.

1. INTRODUCTION

Dynamic visual information from the lip movement can improve the accuracy and robustness of an automatic speech recognition system in a noisy environment [1],[2]. Useful geometric information, such as variation of mouth width and height, can be obtained easily from a segmented lip. However, accurate lip segmentation has problem due to the weak color contrast between lip region and nonlip regions. For color image segmentation, histogram-based and *clustering-based* methods have been widely used.

A histogram segmentation technique which involves performing a fuzzy partition on a two-dimensional (2-D) histogram. Horizontal axis represents one of the pixel color component, whereas the vertical axis represents the local average of that color component, thus capturing the local spatial interactions between neighboring pixels. Color segmentation is obtained by performing the fuzzy partition on each of the color components and then combining the results. Because every the color components are calculated independently, consistent segmentation of a region at every color components is impossible.

A hue filter is used to weight the red hue, which is assumed as the lip region. And then The lip region is segmented by thresholding. This method requires a hue distribution to be pre-specified for the lip region. however, In fact, the lip hue varies between speakers, and is also affected by illumination and by the use of makeup so, it can not give the segmentation result well.

Color lip segmentation is treated as a two-class clustering and segmentation problem Clustering-based method allows the segmentation based on color difference between lip and nonlip regions, without using hue distribution for the lip.

2. METHOD

2.1 Preprocessing Step

In preprocessing step, input image is tranformed into the CIELAB format. The purpose of this process is to get color like eyesight human perception this process produces color feature L*, a*, dan b* that is used to next process. Preprocessing algorithm is shown in Figure 1.

2.2. Lip Membership Computation

This step is segmentation process for lip image color feature that is obtained from preprocessing step will be used to input for segmentation using Spatial Fuzzy Clustering algorithm. segmentation process has two cluster

ISSN 2085-1944

such as lip cluster and skin cluster. Cluster center is initialized randomly.

Before segmentation is done, we need to determine the value (difference of objective function) to stop iteration when the convergen condition is reached. Segmentation using SFCM algorithm flowchart is shown in Figure 2.

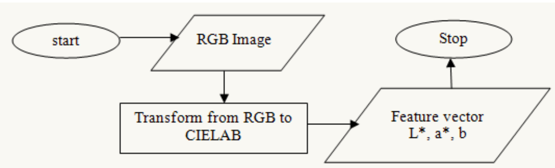

Figure 1. Preprocessing Algorithm Flowchart

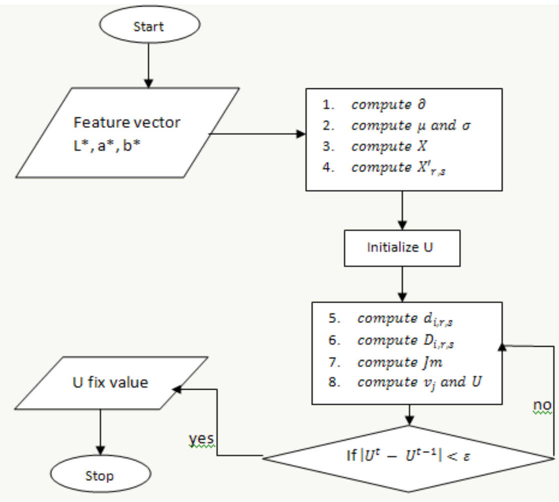

Figure 2. Segmentation using SFCM Algorithm Flowchart

2.2 Post Processing

Post processing is used to smooth the lip membership map, output of segmentation, using SFCM. This process is done in order to get an optimum lip region. There are four sub processes in this post processing.

2.2.1 Morphology

Morphology operation such as: Closing and Opening is done to smooth and omit small hole in the lip membership. This process uses Structure Element disk with size 6, because this structure element produces optimum output. The steps in this process are:

1. Closing is used because the gray value of aimed object is higher than fore ground, therefore this process can fill the small hole.
2. Opening. This process will open small connected region.

2.2.2. Symmetry Processing

The high segmentation error in above and below the lip can not be eliminated using morphology operation. However, the characteristic of lip, symmetry, can be used to solve this problem. Symmetry processing will change all pixels above *y-upper* and under *y-lower* as not lip region by comparing resemblance pixels between left and right image pixel.

First, *xl* (*x* from left corner of lip) and *xr* (*x* from right corner of lip) ,the first five group pixels that has membership map more than 0.5, are achieved by tracking algorithm. Second, integral projection of every line *y* between *xl* and *xc*, and between *xc* and *xr* are obtained using equation (1).

$$\zeta(y) = \sum_{x=xl}^{xc} z(x,y) \qquad (1)$$

where *xc* is *x* from center of lip, *x* coordinate in the center of *xl* and *xr*,

$$z(x,y) = \begin{cases} 1, if\ m(x,y) > 0.5 \\ 0, on\ the\ contrary \end{cases}$$

and *m(x,y)* is lip membership. Third, find coordinate of *y-upper* and *y-lower* that fulfill these conditions in equation 2 and 3.

$$\zeta l(y) > 0 \cap \zeta r(y) > 0 \qquad (2)$$
$$(\zeta l(y) < 3 * \zeta r(y)) \cap (\zeta r(y) < 3 * \zeta l(y)) \qquad (3)$$

The last, change all pixels above *y-upper* and under *y-lower* to be non-lip.

2.2.3. Shape Processing

The special shape of lip, like ellipse, can be used to reduce an inaccurate clustering. A best fit ellipse can be matched with lip membership map to eliminate the rest noises. The parameters for the best fit ellipse are central mass *(xm,xy)*, inclination *(θ)*, semi major axis *(xa)*, and semi minor axis *(ya)*. The pixels computed to get those parameters just pixels that have lip membership more than 0.5, because those pixels have potential to be lip pixel. Those parameters calculated by using equation 4,5,6,7, and 8.

$$xm = \sum_{x=1}^{M} \sum_{y=1}^{N} x * m(x,y) / \sum_{x=1}^{M} \sum_{y=1}^{N} m(x,y) \qquad (4)$$

$$ym = \sum_{x=1}^{M}\sum_{y=1}^{N} y * m(x,y) / \sum_{x=1}^{M}\sum_{y=1}^{N} m(x,y) \quad (5)$$

$$\theta = \frac{1}{2} tan^{-1}\left\{\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right\} \quad (6)$$

$$x_a = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{(I_y)^3}{I_x}\right]^{1/8} \quad (7)$$

$$y_a = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{(I_x)^3}{I_y}\right]^{1/8} \quad (8)$$

$$\mu_{pq} = \sum_{x=1}^{M}\sum_{y=1}^{N}(x - x_m)^p * (y - y_m)^q * m(x,y)$$

$$I_x = \sum_{x=1}^{M}\sum_{y=1}^{N}((y - y_m) * cos\theta - (x - x_m) * sin\theta)^2 * m(x,y)$$

$$I_y = \sum_{x=1}^{M}\sum_{y=1}^{N}((y - y_m) * cos\theta - (x - x_m) * sin\theta)^2 * m(x,y)$$

After those parameters of best ellipse are found, draw ellipse. Then, get the intersection between lip membership function and ellipse using AND operation. Noises are reduced by changing all pixels outside the best fit lip to non lip pixel.

2.2.4. The Last Processing

In the last processing, the lip memberships from shape processing are smoothed. The smoothing processes are:

1. Eliminate the rest hole in the center of the lip.
2. Get single connected component.
3. Draw a line in the lip region.

3. MEASURING EVALUATION

From the testing of ROI (Region Of Interest) on colored lip image, the accuracy of this algorithm is measured. The segmented image from this application is compared with the manually segmented image. Error is evaluated using SE (Square Error) method defined in the equation bellow:

$$SE = \frac{OLE + ILE}{TL} * 100\%$$

where OLE (Outer Lip Error) is nonlip pixel classified as lip pixel, ILE (Inner Lip Error) is lip pixel classified as nonlip pixel, and TL is the number of pixel in image.

4. RESULT

The testing consists of size neighborhood test and difference among objective functions. The input of this application are color lip images from COLORFERET database [3]. In the first test, image will be tested with neighborhood size 3x3, 5x5, dan 7x7.

In the second test, image will be tested with value of threshold from 0 to 1. The implementation of application uses Matlab 7.0.

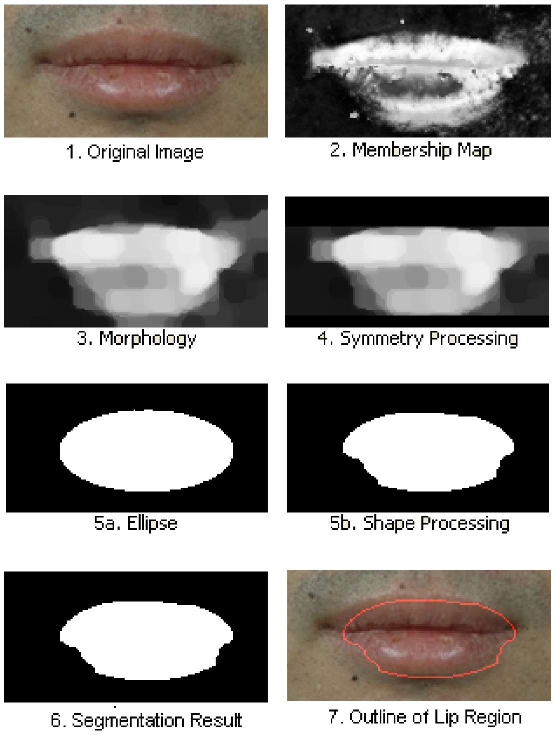

Figure 3 Experimental Image

In the experiment is used image 029_19.ppm with size neighborhood 3x3, maximum iteration 500, and difference objective functions threshold 0.001. The run time for this experiment is 8.11 second. The iteration has stopped in 74^{th} iteration.

The experimental result was compared with the ground truth and the accuracy is 92.74%. Figure 3.2 shows the image result of each process. Figure 3.2(1) shows input image, Figure 3.2(2) shows the lip membership from segmentation result using SFCM.

Morphology operation is done to smooth noise that shown in Figure 3.2(3). The next step is noise removal at top and bottom of the object using symmetry processing that shown in Figure 3.2(4). Figure 3.2(5a) is ellipse shape that part of shape processing. The parameters of ellipse shape are obtained from lip membership of symmetry processing result. Ellipse is used to determine the lip region. If the coordinates of lip membership are in outside of the lip region, the pixels that are in its coordinates are changed to nonlip region. Figure 3.2(5b) is shape processing output.

Figure 3.2(6) is the smoothing result from Figure 3.2(5b) in the last processing. Smoothing in this experiment is done to remove hole in object center and find the single connected component. Figure 3.2(7) is outline of lip region in original image.

4.1 Evaluation of Neighborhood Size

Table 1 shows the evaluation of neighborhood test for every tested image. The shade columns show the size of neighborhood producing the best accuracy and need the least running time. The average accuracy of neighborhood size shows that the highest accuracy is reached when the size of neighborhood is 3x3, and the lowest accuracy is reached when the neighborhood size is 7x7 with value of accuracy is 93,95% and 93,04% respectively. The use of big neighborhood producing low accuracy. Therefore the size of neighborhood is opposite with the value of accuracy. The average of running time for segmenting the image reached by the neighborhood with size 3x3 and 7x7 are 6.76 seconds and 14.3 seconds, respectively.

4.2 Evaluation Of The Objective Function Difference

The result of difference among objective functions threshold is shown in Table 2. Using image test 091.ppm the stable accuracy is achieved when the threshold is 0.5.

Table 1. Evaluation Of Neighborhood Size

No	Input Image	Neighborhood Size					
		3x3		5x5		7x7	
		Running Time (second)	Accuracy(%)	Running Time (second)	Accuracy (%)	Running Time (second)	Accuracy(%)
1	029_19.ppm	8.11	92.74	10.02	91.89	18.84	91.7
2	071.ppm	5.41	95.15	7.67	94.5	9.72	94.38
	Average	6.76	93.95	8.85	93.2	14.3	93.04

While when using image 093.ppm and 130.ppm, the stable accuracy is achieved when the thresholds are 0.4 and 0.6. From the three experiments, the optimum value of threshold is 0.6.

Table 2. Evaluation Of Threshold Value

No	Input image	Threshold
1	091.ppm	0.5
2	093.ppm	0.4
3	130.ppm	0.6
Maximum		0.6

5. CONCLUSION

Based on the testing and evaluation process, can be concluded that:

a. The addition of local spatial interaction among neighborhood pixels is effective on segmentation.
b. The smoothing process produces high accuracy in segmentation.
c. The size of neighborhood influences accuracy. The bigger the size of neighborhood, the lower the accuracy because the distance of the pixels are larger. From the testing process, the optimum neighborhood size is 3x3.
d. The size of neighborhood influences the run time. The bigger the neighborhood's size the longer the run time.
e. The different of the threshold value between two objective functions influence the accuracy of the segmentation. From the testing process the optimum value of threshold is 0.6.
f. The different of illumination among the image pixels causes the segmentation is not optimum.

REFERENCES

[1]. A.Liew, S.H.Leung, dan W.H.Lau. "Segmentation of Color Lip Images by Spatial Fuzzy Clustering". IEEE Transactions on Fuzzy Systems, 2003, 11:542–549.

[2]. A. W. C. Liew, S. H. Leung, dan W. H. Lau, "Fuzzy Image Clustering Incorporating Spatial Continuity". Proc. Inst. Elect. Eng.—Vision, Image, Signal Processing, vol. 147, no. 2, Apr. 2000, hal. 185–192.

[3]. Database Citra Wajah Berwarna COLORFERET [Online]. http://face.nist.gov/colorferet.

[4]. H. D. Cheng, J. R. Chen, dan J. Li, "Threshold Selection Based on Fuzzy C-Partition Entropy *Approach"*, Pattern Recogn., vol. 31, no. 7, 1998, hal. 857–870.

[5]. H. D. Cheng, Y. H. Chen, dan X. H. Jiang, "*Thresholding Using Two-Dimensional Histogram And Fuzzy Entropy Principle*".

IEEE Trans. Image Processing, vol. 9, Apr. 2000, hal. 732–735.

[6]. T. Coianiz, L. Torresani, dan B. Caprile, "*2D Deformable Models for Visual Speech Analysis*". Speechreading by Humans and Machines, D. G. Stork and M. E. Hennecke, Eds. New York: Springer-Verlag,1996.

[7]. Tutorial *Fuzzy C-Means Clustering* [Online]. http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html.

[This page intentionally left blank]

ISSN 2085-1944