

Automatic Corneal Ulcer Segmentation Combining Gaussian Mixture Modeling and Otsu Method

Zhenrong Liu^{1,2}, Yankun Shi^{1,2}, Pengji Zhan^{1,2}, Yue Zhang^{1,3}, Yi Gong^{1,*}, Xiaoying Tang^{1,*}

Abstract—In this paper, we proposed and validated a novel and accurate pipeline for automatically segmenting flaky corneal ulcer areas from fluorescein staining images. The ulcer area was segmented within the cornea by employing a joint method of Otsu and Gaussian Mixture Modeling (GMM). In the GMM based segmentation, the total number of Gaussians was determined intelligently using an information theory based algorithm. And the fluorescein staining images were processed in the HSV color model rather than the original RGB color model, aiming to improve the segmentation results' robustness and accuracy. In the Otsu based segmentation, the images were processed in the grayscale space with Gamma correction being conducted before the Otsu binarization. Afterwards, morphological operations and median filtering were employed to further improve the Otsu segmentation result. The GMM and Otsu segmentation results were then intersected, for which post-processing was conducted by identifying and filling holes through a fast algorithm using priority queues of pixels. The proposed pipeline has been validated on a total of 150 clinical images. Accurate ulcer segmentation results have been obtained, with the mean Dice Similarity Coefficient (DSC) being 0.88 when comparing the automatic segmentation result with the manually-delineated gold standard. For images in the RGB color space, the mean DSC was 0.83, being much lower than that of the images in the HSV color space.

Index Terms—Ocular staining images, Corneal ulcers, Image segmentation, Ostu, Gaussian mixture model, HSV

I. INTRODUCTION

A corneal ulcer is an inflammatory or more seriously infective condition of the cornea involving disruptions of its epithelial layer with involvement of the corneal stroma, which brings huge pain and may cause serious loss of vision or even blindness [1-2]. Fluorescein is a common diagnostic dye used to assess the ocular surface integrity because of its fluorescent properties and its high visibility at low concentrations. When assessing a corneal ulcer using fluorescein staining, the ulcer area appears luminous green whereas the rest areas of the cornea appear blue or brown [3]. The morphometry and chromaticity of ulcer areas are important criteria for doctors to assess the severity of a patient's corneal ulcer for which a prerequisite is to segment out the ulcer area.

This study was supported by the National Key R&D Program of China (2017YFC0112404) and the National Natural Science Foundation of China (NSFC 81501546).

*Correspondence to Dr. Xiaoying Tang: tangxy@sustc.edu.cn and Prof. Yi Gong: gongy@sustc.edu.cn

¹Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China

²Peng Cheng Laboratory, Shenzhen, China

³Laboratory of Biomedical Imaging and Signal Processing and Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China

Usually, professional softwares such as Photoshop or ImageJ are used to perform ulcer area extraction manually [4]. Despite the high accuracy, it is extremely tedious, time consuming and prone to subjective errors. Automatic extraction methods with high accuracies have always been desired. In the past decade, there have been many semi- and fully-automatic approaches for ophthalmic image segmentation. Welfer *et al.* [5] used a morphological approach to segment the optic disk from color fundus images. Ferreira *et al.* [6] proposed an automatic method for extracting corneal nerves. Arsalan *et al.* [7] performed automatic iris segmentation using deep learning. However, researches on automatic corneal ulcer segmentation are relatively rare, especially for flaky ulcers. In one of our previous works, we attempted semi-automatic ulcer segmentation utilizing k -means clustering followed by morphological operations and region growing [8].

The Otsu method and Gaussian Mixture Modeling (GMM) have been widely used in image segmentation tasks [9]. Both methods have been proven to be useful in biological studies [10-11]. Compared to the recently-developed deep learning based segmentation techniques, the main advantage of those classical approaches is they are unsupervised and thus no need for manually-created training data.

In this paper, we propose a novel method for automatically segmenting flaky corneal ulcers from fluorescein staining images, combining Otsu thresholding and GMM. The color images are converted from the original RGB (Red, Green, Blue) space to the HSV (Hue, Saturation, Value) space. HSV is a color space representing the intuitive nature of color [12]. The expectation-maximization (EM) algorithm is used to automatically determine the parameters in the proposed GMM. Gamma correction is conducted with respect to the grayscale image before the Otsu based thresholding, after which morphological operations and median filtering are performed. The proposed method is quantitatively validated using the Dice Similarity Coefficient (DSC) between the ground truth segmentation and the automatic segmentation [13]. In addition, we investigate the potential underlying reason why HSV-based segmentation outperforms RGB-based by comparing the within-class and between-class distances in the GMM setting.

II. METHOD

The proposed corneal ulcer segmentation pipeline combines Otsu thresholding and GMM, as shown in Fig. 1.

A. Otsu Thresholding in Grayscale Space

In Otsu thresholding, Gamma correction is conducted after transforming the original RGB images into grayscale images. Gamma correction makes the contrast between staining areas and background larger. The value γ in the Gamma correction is empirically set to be 0.9 in $I_{out} = I_{in}^\gamma$, $\gamma < 1$, where I_{in} denotes the intensity of the grayscale input image and I_{out} denotes the corresponding image intensity after Gamma correction. After that, we use Otsu thresholding for image binarization. For a better segmentation result, morphological operations (opening followed by closing with a radius of 3) and median filtering are employed afterwards.

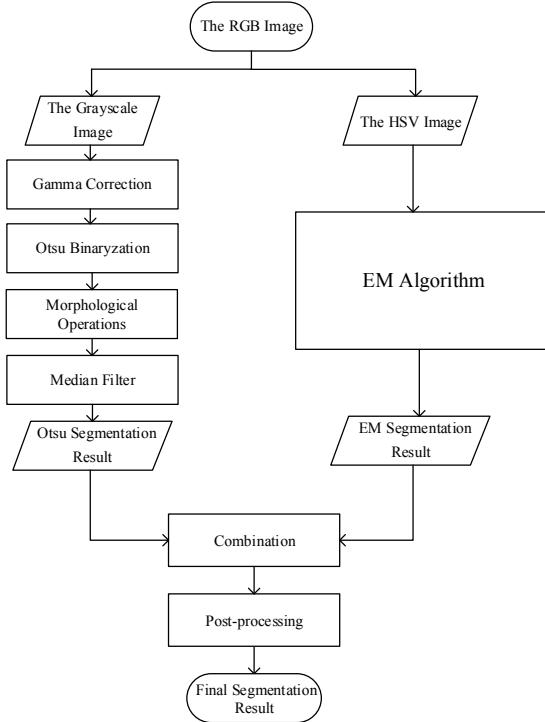


Fig. 1. The entire flow chart of the proposed ulcer segmentation pipeline.

B. GMM in HSV Space

An important issue in GMM is the selection of the total number of Gaussian components K . The deterministic methods start with a set of candidate models for different values of K which are supposed to contain the true/optimal K . The optimal value is then selected according to some model selection criterion. Usually, the model selection criterion based on information/coding theory has the following form [14-15]:

$$\mathcal{C}(\hat{\theta}(K), K) = -\log p(\mathcal{Y}|\hat{\theta}(K)) + \mathcal{P}(K), \quad (1)$$

where $\mathcal{Y} = \{y_1, y_2, \dots, y_N\}$ denotes a set of identically distributed samples, $\hat{\theta}(K)$ denotes a complete set of estimated parameters needed to specify the GMM with K components, $-\log p(\mathcal{Y}|\hat{\theta}(K))$ denotes the code length of the data samples, $\mathcal{P}(K)$ denotes an increasing function penalizing higher values of K .

In our work, we use the model selection criterion proposed by Figueiredo *et al.* in [16]:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}, \mathcal{Y}) = & \frac{P}{2} \sum_{k:\pi_k>0} \log\left(\frac{N\pi_k}{12}\right) + \frac{K}{2} \log\left(\frac{N}{12}\right) \\ & + \frac{K(P+1)}{2} - \log p(\mathcal{Y}|\hat{\theta}(K)), \end{aligned} \quad (2)$$

where P is the number of parameters specifying each component, N is the number of data samples (pixels in our case), π_k is the k th mixing coefficient and K is the number of components.

Please note, the optimal value of K may vary from image to image. In our experiments, 2 is the most frequently occurred value of K . As such, in our work, the value of K is set to be 2.

After converting a to-be-segmented image from RGB to HSV, we build a three-dimensional Gaussian Mixture Model with two components:

$$p(\mathbf{x}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{k=1}^2 \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (3)$$

where $\mathbf{x} = [h, s, v]^T$ and h, s, v respectively denote the image value of each channel in HSV, $\boldsymbol{\mu} = [\mu_h, \mu_s, \mu_v]^T$ and μ_h, μ_s, μ_v respectively denote the mean value of h, s, v , $\boldsymbol{\Sigma}$ denotes the covariance matrix of Gaussian components.

Given an image with N pixels $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, the log-likelihood of all pixels is given by:

$$\begin{aligned} \ln p(\mathbf{X}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \ln \prod_{n=1}^N p(\mathbf{x}_n|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \sum_{n=1}^N \ln \sum_{k=1}^2 \pi_k \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k). \end{aligned} \quad (4)$$

To maximize the log-likelihood given in Eq. (4), the EM algorithm is applied to estimate the parameters $\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$ as below:

- Initialization:

Use k -means ($k = 2$) clustering to obtain the initial values of $\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$.

- E step:

$$\gamma_{nk} = \frac{\pi_k \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^2 \pi_j \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}, \quad (5)$$

where γ_{nk} denotes a posterior probability that pixel n comes from component k under the current model parameters, which is considered as the response of component k to pixel n .

- M step:

$$\boldsymbol{\mu}_k^{new} = \frac{\sum_{n=1}^N \gamma_{nk} \mathbf{x}_n}{\sum_{n=1}^N \gamma_{nk}}, \quad (6)$$

$$\Sigma_k^{new} = \frac{\sum_{n=1}^N \gamma_{nk} (\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^T}{\sum_{n=1}^N \gamma_{nk}}, \quad (7)$$

$$\pi_k^{new} = \frac{\sum_{n=1}^N \gamma_{nk}}{N}. \quad (8)$$

- Repeat E step and M step until:

$$\ln p(\mathbf{X}|\boldsymbol{\pi}^{new}, \boldsymbol{\mu}^{new}, \boldsymbol{\Sigma}^{new}) - \ln p(\mathbf{X}|\boldsymbol{\pi}^{old}, \boldsymbol{\mu}^{old}, \boldsymbol{\Sigma}^{old}) < \epsilon. \quad (9)$$

In our experiment, the stopping criterion ϵ is set to be 10^{-5} . Once EM finishes, we perform component matching for each pixel to achieve segmentation. We summarize the entire GMM based segmentation in Algorithm 1.

Algorithm 1 GMM Based Segmentation

```

1: Input: HSV image  $\mathbf{X}$  (with  $N$  pixels), component number  $K$ , decision threshold  $\epsilon$ 
2: Initialization:  $t \leftarrow 0$ , mixing coefficients  $\pi_k$ , means  $\boldsymbol{\mu}_k$ , and covariances  $\boldsymbol{\Sigma}_k$ ,  $k = 1, 2$ 
3: Calculate log-likelihood  $l(t) = \ln p(\mathbf{X}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ 
4: do
5:    $t \leftarrow t + 1$ 
6:   E step: Compute  $\gamma_{nk}$  using current  $\pi_k$ ,  $\boldsymbol{\mu}_k$  and  $\boldsymbol{\Sigma}_k$ 
7:   M step: Update  $\pi_k^{new}$ ,  $\boldsymbol{\mu}_k^{new}$  and  $\boldsymbol{\Sigma}_k^{new}$ 
8:   Recalculate log-likelihood as  $l(t)$ 
9: while  $l(t) - l(t - 1) > \epsilon$ 
10: Initial a binary image  $\mathbf{X}_{EM}$  (with  $N$  pixels)
11: for  $n = 1, 2, \dots, N$  do
12:   if  $\gamma_{n1} > \gamma_{n2}$  then
13:     Set the value of pixel  $n$  from  $\mathbf{X}_{EM}$  to 0
14:   else
15:     Set the value of pixel  $n$  from  $\mathbf{X}_{EM}$  to 1
16:   end if
17: end for
18: Output: The segmentation result  $\mathbf{X}_{EM}$ 

```

C. Post-processing

After obtaining the results of Otsu segmentation and EM segmentation, we combine them by taking intersection. Certain segmentation results exist abnormal holes due to the reflection in the images, which may have been induced from uneven lighting. In morphological reconstruction, holes within a binary image correspond to a set of regional minima that are not connected to the area border [17]. In our work, we identify and fill the holes through a fast algorithm using priority queues of pixels [18].

D. Evaluation

We use DSC between an automated segmentation and the corresponding ground truth segmentation to quantitatively assess the performance of the proposed pipeline, which is defined as:

$$DSC(S, G) = \frac{2|S \cap G|}{|S + G|}, \quad (10)$$

where S and G respectively denote the automatic segmentation result and the corresponding ground truth, $|a|$ denotes the area size of image a .

III. EXPERIMENTS AND RESULTS

A. Dataset

The dataset used in this study, obtained from 150 patients with flaky corneal ulcers of different degrees of severity, were collected at Zhong-shan Ophthalmic Centre, Sun Yat-sen University. Photographs of the ocular surfaces were taken with the eye in the middle position after fluorescein instillation, and were selected on the basis of high image quality. Each photograph was interfaced to a personal computer and saved as a JPG file (2592×1728 pixels, 2 MB, RGB). For validation purpose, the corneal ulcer area has been manually delineated from each staining image using the ImageJ software.

B. Segmentation Results

Upon visual examination, the proposed pipeline was able to accurately segment out the corneal ulcers of all 150 images. In Fig. 2, we show the comparison results among the ground truth, the proposed pipeline in RGB and HSV. Apparently, the automatic segmentation results obtained from the proposed pipeline are very close to the gold standard. The corresponding means and standard deviations of DSC calculated across all 150 images are tabulated in TABLE I. Evidently, a combination of Otsu thresholding and GMM with application to images in the HSV space yielded the best segmentation performance.

TABLE I
MEANS AND STANDARD DEVIATIONS OF DSC OBTAINED
FROM DIFFERENT SEGMENTATION SETTINGS.

	Otsu [%]	GMM [%]	GMM+Otsu [%]
RGB	57.49 ± 20.35	81.69 ± 11.67	83.32 ± 10.20
HSV		86.59 ± 7.83	88.05 ± 6.11

C. Clustering Analysis

In addition to comparing the difference between RGB and HSV based segmentation performance, we also explore the potential underlying reason why HSV-based segmentation performs better than RGB-based through clustering analysis. GMM based segmentation can be treated as a clustering problem. Using each cluster's mean and variance computed from EM, we can compare the clustering performance of the two color models. The between-class distance and within-class distance are typically used to evaluate a clustering algorithm. In our case, the between-class distance is computed as the Euclidean distance between two Gaussian distributions' mean and the within-class distance is computed as the trace

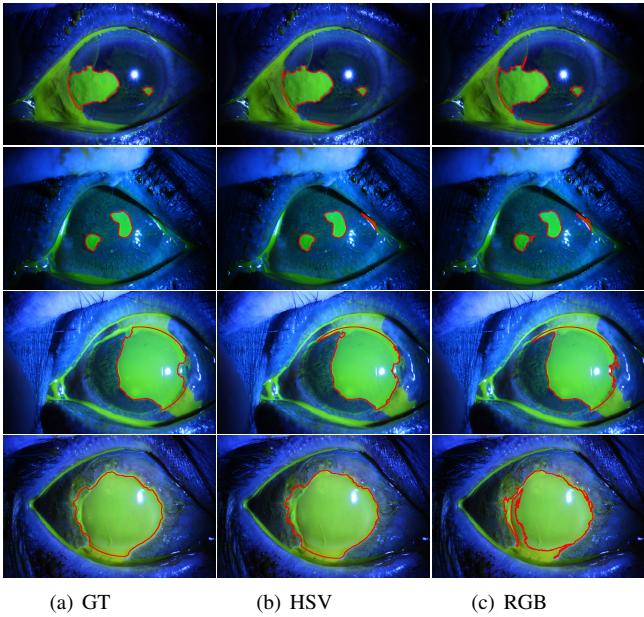


Fig. 2. A comparison of results obtained from manual delineation (first column), the proposed pipeline in HSV color model (middle column), and the proposed pipeline in RGB color model (last column). Please note, the green color represents Fluorescein dye.

of the corresponding covariance matrix. As shown in TABLE II, the between-class distance of HSV is slightly larger than that of RGB. However, the within-class distance of HSV is significantly smaller than that of RGB, being about half of that of RGB, especially for the ulcer area. This clustering analysis indicates that the HSV color model is superior to the RGB color model in representing the corneal ulcer features.

TABLE II

WITH-IN CLASS DISTANCE AND BETWEEN-CLASS DISTANCE OF HSV AND RGB. "U" AND "B" RESPECTIVELY DENOTE THE ULCER CLASS AND THE BACKGROUND CLASS.

	With-in class distance (U)	With-in class distance (B)	Between-class distance
RGB	0.056 ± 0.031	0.086 ± 0.036	0.487 ± 0.219
HSV	0.021 ± 0.014	0.052 ± 0.016	0.530 ± 0.097

IV. CONCLUSION

In this work, we proposed and validated a pipeline for automatically extracting flaky corneal ulcers by combining Otsu thresholding and GMM fitting. Based on a total of 150 fluorescein staining images, we have validated the superiority of the proposed method in terms of segmentation accuracy. Our experiment results indicate that the HSV color model is better than RGB in terms of corneal ulcer segmentation. The encouraging results observed from this work suggest that the proposed pipeline may be useful for an objective assessment of corneal ulcers. It also has a great potential in monitoring corneal health. With that being said, there are still limitations. First of all, in this pipeline, the convergence threshold used

in the EM algorithm was determined empirically, which may not be the optimal one. An automatic and intelligent approach to identify that value is desired. Another issue is that the proposed pipeline has been only validated on flaky corneal ulcers. Future work will also involve segmenting corneal ulcers of other types of patterns such as point-like ulcers. Meanwhile, the techniques used in this pipeline are mostly traditional image processing techniques. A combination of classical image processing techniques and recently-developed deep learning techniques may further boost our segmentation performance, which will also be one of our future research directions.

REFERENCES

- [1] Smith T. BMA AZ Family Medical Encyclopedia[M]. Dorling Kindersley Ltd, 2004.
- [2] Deswal J, Arya S K, Raj A, et al. A Case of Bilateral Corneal Perforation in a Patient with Severe Dry Eye[J]. Journal of clinical and diagnostic research: JCDR, 2017, 11(4): ND01.
- [3] Morgan P B, Maldonado-Codina C. Corneal staining: do we really understand what we are seeing?[J]. Contact Lens and Anterior Eye, 2009, 32(2): 48-54.
- [4] Tang X N, Berman A E, Swanson R A, et al. Digitally quantifying cerebral hemorrhage using Photoshop® and Image J[J]. Journal of neuroscience methods, 2010, 190(2): 240-243.
- [5] Welfer D, Scharcanski J, Kitamura C M, et al. Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach[J]. Computers in Biology and Medicine, 2010, 40(2): 124-137.
- [6] Ferreira A, Morgado A M, Silva J S. A method for corneal nerves automatic segmentation and morphometric analysis[J]. Computer methods and programs in biomedicine, 2012, 107(1): 53-60.
- [7] Arsalan M, Hong H G, Naqvi R A, et al. Deep learning-based iris segmentation for iris recognition in visible light environment[J]. Symmetry, 2017, 9(11): 263.
- [8] Deng L, Huang H, Yuan J, et al. Automatic segmentation of corneal ulcer area based on ocular staining images[C]/Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging. International Society for Optics and Photonics, 2018, 10578: 105781D.
- [9] Otsu N. A threshold selection method from gray-level histograms[J]. IEEE transactions on systems, man, and cybernetics, 1979, 9(1): 62-66.
- [10] Song Y, Ji Z, Sun Q. An extension Gaussian mixture model for brain MRI segmentation[C]/2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2014: 4711-4714.
- [11] Jeong H J, Kim T Y, Hwang H G, et al. Comparison of thresholding methods for breast tumor cell segmentation[C]/Enterprise networking and Computing in Healthcare Industry, 2005. HEALTHCOM 2005. Proceedings of 7th International Workshop on. IEEE, 2005: 392-395.
- [12] Smith A R. Color gamut transform pairs[J]. ACM Siggraph Computer Graphics, 1978, 12(3): 12-19.
- [13] Dice L R. Measures of the amount of ecological association between species[J]. Ecology, 1945, 26(3): 297-302.
- [14] Oliver J J, Baxter R A, Wallace C S. Unsupervised learning using MML[C]/ICML. 1996: 364-372.
- [15] Wallace C S, Dowd D L. Minimum message length and Kolmogorov complexity[J]. The Computer Journal, 1999, 42(4): 270-283.
- [16] Figueiredo M A T, Jain A K. Unsupervised learning of finite mixture models[J]. IEEE Transactions on pattern analysis and machine intelligence, 2002, 24(3): 381-396.
- [17] Soille P. Morphological image analysis: principles and applications[M]. Springer Science & Business Media, 2013.
- [18] Soille P, Gratin C. An Efficient Algorithm for Drainage Network Extraction on DEMs[J]. Journal of Visual Communication and Image Representation, 1994, 5(2):181-189.