Toward real-time quantification of fluorescence molecular probes using target/background ratio for guiding biopsy and endoscopic therapy of esophageal neoplasia

Yang Jiang
Yuanzheng Gong
Joel H. Rubenstein
Thomas D. Wang
Eric J. Seibel

Toward real-time quantification of fluorescence molecular probes using target/background ratio for guiding biopsy and endoscopic therapy of esophageal neoplasia

Yang Jiang,^a Yuanzheng Gong,^b Joel H. Rubenstein,^{c,d} Thomas D. Wang,^c and Eric J. Seibel^{b,*}

^aUniversity of Washington, Department of Bioengineering, Human Photonics Lab, Seattle, Washington, United States

^bUniversity of Washington, Department of Mechanical Engineering, Human Photonics Lab, Seattle, Washington, United States

^cUniversity of Michigan, Division of Gastroenterology, Department of Internal Medicine, Ann Arbor, Michigan, United States

^dVeterans Affairs Center for Clinical Management Research, Ann Arbor, Michigan, United States

Abstract. Multimodal endoscopy using fluorescence molecular probes is a promising method of surveying the entire esophagus to detect cancer progression. Using the fluorescence ratio of a target compared to a surrounding background, a quantitative value is diagnostic for progression from Barrett's esophagus to high-grade dysplasia (HGD) and esophageal adenocarcinoma (EAC). However, current quantification of fluorescent images is done only after the endoscopic procedure. We developed a Chan–Vese-based algorithm to segment fluorescence targets, and subsequent morphological operations to generate background, thus calculating target/background (T/B) ratios, potentially to provide real-time guidance for biopsy and endoscopic therapy. With an initial processing speed of 2 fps and by calculating the T/B ratio for each frame, our method provides quasireal-time quantification of the molecular probe labeling to the endoscopist. Furthermore, an automatic computer-aided diagnosis algorithm can be applied to the recorded endoscopic video, and the overall T/B ratio is calculated for each patient. The receiver operating characteristic curve was employed to determine the threshold for classification of HGD/EAC using leave-one-out cross-validation. With 92% sensitivity and 75% specificity to classify HGD/EAC, our automatic algorithm shows promising results for a surveillance procedure to help manage esophageal cancer and other cancers inspected by endoscopy. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.4.2.024502]

Keywords: target/background ratio; surgical guidance; multimodal endoscopy; esophageal cancer detection; computer-aided diagnosis; fluorescence video processing.

Paper 17044PR received Feb. 17, 2017; accepted for publication Apr. 24, 2017; published online May 24, 2017.

1 Introduction

Esophageal adenocarcinoma (EAC) possesses a 5-year survival rate less than 18% in the US with increasing incidence.^{1,2} Barrett's esophagus (BE) is regarded as a premalignant condition of EAC. BE is defined as the metaplastic change of the esophagus, in which a normal squamous (SQ) epithelium has been transformed to a columnar epithelium containing goblet cells.³ The transformations from BE to EAC progress through low-grade dysplasia (LGD) and high-grade dysplasia (HGD).⁴ Early detection and treatment of EAC at its premalignant stages could increase the survival rate of the patients. Clinically, endoscopic surveillance is often recommended to BE patients. Current standard protocol is applying high-resolution white-light endoscopy with four-quadrant biopsies at 1- to 2-cm intervals.⁵ However, this approach is intensive and expensive. Due to the lack of visualization of surface abnormality present in dysplastic lesions, extensive biopsies are randomly taken. The histopathological processing of these biopsies provides slow feedback to the endoscopist, often 1 to 2 weeks. Since only a small fraction of the surface area is sampled (4% to 6%), it is still an inaccurate evaluation of dysplasia or EAC due to the sampling error.⁶ Therefore, advanced imaging for more

comprehensive biopsy guidance and more rapid and accurate detection of dysplastic tissue progressing to EAC is highly demanded.

Several advanced endoscopic imaging techniques have been investigated to detect dysplasia associated with BE.^{7–9} Autofluorescence imaging (AFI) is a phenomenon where endogenous substances, such as collagen and flavins, emit longer wavelength light after being excited with short wavelength light. By distinguishing among different autofluorescence patterns highlighted by AFI, an experienced endoscopist can find more dysplastic lesions from the background of normal SQ epithelium. Boerwinkel et al.¹⁰ showed that AFI possesses an 89% sensitivity of HGD and EAC, yet suffers from high false positive rates of 86%. This is contradictory to their previous findings concluding that an earlier commercial version of AFI had limited ability in detecting dysplasia lesions.⁷ Narrow band imaging (NBI) illuminates the mucosa in blue band using laser light around 415 nm, which targets hemoglobin absorption, to enhance the visualization of vascular structures at the mucosal surface.¹¹ Wolfsen et al.⁸ showed NBI can identify patients with dysplasia with significantly (50%) fewer biopsy samples. Recent studies showed that detailed inspection with NBI after

*Address all correspondence to: Eric J. Seibel, E-mail: eseibel@uw.edu

AFI reduced the false positive rate from 81% to 26%.¹² Overall, both NBI and AFI use nonspecific contrast mechanisms to detect dysplasia lesions, which suffer from low specificity. Angle-resolved low coherence interferometry, a type of optical biopsy, measures the nuclear morphology changes to evaluate dysplastic changes in the esophageal mucosa. However, this technique relies on point sampling, which is not a method for interrogating the entire mucosa having BE.¹³ Volumetric laser endomicroscopy and tethered capsule endomicroscopy are comprehensive imaging techniques, which can provide frequency-domain optical coherence tomography (OCT) imaging from a minimally invasive side-viewing scope.^{9,14} However, they can only provide cross-sectional structural data, without molecular or chemical contrast, which may be required for detecting cells progressing from LGD to HGD.

Molecular imaging is an emerging technique that can provide molecular biology information specific to cancer progression, which is showing promise in the management of gastrointestinal (GI) cancers.^{15,16} Optical imaging techniques, such as wide-field endoscopy or smaller field of view endomicroscopy, are used to measure relative intensity of fluorescently labeled probes that are associated with EAC. Using a fluorescence agent that indicates high metabolic activity of neoplastic EAC *ex vivo*, Leggett et al.¹⁷ achieved a sensitivity of 79% and a specificity of 77% using probe-confocal laser endomicroscopy (pCLE). However, the fluorescent marker they applied, 2-NBDG, is a nontargeted fluorescence agent and limited by nonspecific binding. Bird-Lieberman et al.¹⁸ used fluorescence endoscopy with fluorescently labeled lectin to detect the changes of glycan following the progression from BE to EAC *in situ*. However, this study was limited by the small number of patient tissue samples and failed to report the sensitivity and specificity of the proposed technique in targeting dysplastic lesions compared to the state of the art. Moreover, lectin is a negative contrast agent, which is limited by false positives due to low specific binding *in vivo* because of overlying mucus and mucosal folds. By discovering a fluorescent peptide that specifically binds to EAC neoplasia, Hsiung et al.¹⁶ attained 75% sensitivity and 97% specificity with a first-in-human study using pCLE imaging and histological correlation. By using the same molecular specific fluorescence peptide as Hsiung et al.¹⁶ and including more subjects (50 subjects) in the study, Joshi et al.¹⁹ achieved 76% sensitivity and 94% after postprocessing the fluorescence video frames from a wide-field custom fluorescence endoscope. Joshi et al. assessed the feasibility of using fluorescence-labeled target intensity compared to the background intensity to localize HGD and EAC with histological correlation for computer-aided diagnosis (CAD). However, a major limitation was the slow image processing that was done after the endoscopic procedure to calculate diagnostic target/background (T/B) ratios. This computational limitation would not allow the endoscopist to use T/B analysis as an indicator during the endoscopic procedure to guide biopsy or resection. Real-time quantification and mapping of the fluorescence image are required as a navigation guide for sampling the highest-risk regions for dysplasia, EAC, as well as providing direct feedback to surgeons when removing these areas and neoplasia by endoscopic mucosal resection (EMR). Moreover, in future applications of surveillance in rural clinics, automatic interpretation of fluorescence images is strongly desired to assist in making diagnostic decisions. Qi et al.²⁰ achieved automatic interpretation of endoscopic OCT images to detect dysplasia

in BE with a sensitivity of 82% and a specificity of 74%. However, to the best of our knowledge, there is no image processing algorithm for wide-field fluorescence endoscopy to (1) guide the endoscopist to the most suspicious regions in BE during interventions of biopsy and EMR in real time and (2) to automatically analyze multimodal videos to classify these regions as high-risk dysplastic or neoplastic lesions.

One of the main challenges to achieve (1) real-time guidance and (2) automatic CAD of BE patients using fluorescence endoscopy is to select an accurate and fast algorithmic approach to label the foreground targets in each frame of the fluorescence video. Due to the low-contrast “patchy” appearance of the fluorescence targets and nonuniform intensity in both targets and background, the segmentation algorithms using a single threshold value, such as Otsu’s algorithm, show unsatisfactory results. These limitations were demonstrated by Chan et al.²¹ when applying to fluorescence images. Otsu’s algorithm assumes a bimodal intensity histogram of images, while the intensity histograms of fluorescent images are normally unimodal due to large fluorescent signal variation. When applying a single threshold to these unimodal-histogram fluorescent images, small targets with lower intensity will be mistakenly excluded from the labeled target region. On the other hand, the Chan–Vese algorithm²² has the ability to adapt the boundary of the targets iteratively and locally, thus this algorithm is more robust in noisy low-contrast images. Moreover, efficient processing of the Chan–Vese algorithm allows for real-time display of measured and highlighted regions. To our knowledge, this algorithm has never been applied on fluorescence endoscopic video.

In this paper, we describe an approach to achieve highly efficient calculation of T/B ratio from fluorescence videos and demonstrate a method of highlighting regions of interest (ROIs) that score above our T/B threshold for direct feedback to the endoscope user (part 1). Part of this method was presented in Multimodal Biomedical Imaging panel within 2017 SPIE Photonics West BiOS.²³ We expand the results and discussion in this journal. In addition, we develop an automatic CAD algorithm for classifying EAC/HGD from recorded fluorescence videos of esophageal endoscopy with molecular probes (part 2). In recorded videos from 50 subjects, we show that a surveillance protocol based on this automated algorithm using fluorescently labeled peptides specific for BE neoplasia is capable of classifying precancerous and neoplastic (HGD and EAC) regions with 92% sensitivity and 75% specificity using histological confirmation.

2 Materials and Method

2.1 Patients and Endoscopic Procedure

The primary data for the analyses were described by Joshi et al.¹⁹ Fifty patients with a history of BE participated in the diagnosis of HGD/EAC and the endoscopic treatment took place at the University of Michigan. Fluorescence and reflectance videos were collected using a customized Olympus endoscope (GIF-Y0029) having 20 frames per second and 140-deg maximum light collection angle to the camera. Peptide ASY-NYDA, specific binding to early Barrett’s neoplasia, is labeled with fluorescein isothiocyanate (FITC), hereafter called ASY*-FITC.¹⁶ Topical (sprayed) ASY*-FITC peptide was administered. After a short incubation time (5 min) and rinsing to remove unbound peptides, a time-series of white-light, fluorescence, and reflectance videos were collected on selected

segments of BE. After clinical evaluation, biopsies and EMR tissue specimens were removed from the same BE segments. The tissue was sectioned in 2-mm intervals, and a diagnosis was determined by two gastrointestinal pathologists. Disease classification was recorded as positive (HGD or EAC) or negative [SQ, BE, gastroesophageal junction (GEJ), or LGD] based on histological interpretation, see Table 2 within Joshi et al.¹⁹

2.2 Part 1: Target/Background Ratio Calculation

T/B ratio was calculated for each frame in the recorded wide-field fluorescence videos. All these videos were recorded in high resolution (1920×1080) with 8-bit intensity using an Olympus image digitizer. The overall procedure to calculate T/B ratio is shown in Fig. 1.

For each frame, an image-smoothing step is first performed using a Gaussian filter with 5×5 window. After smoothing, ROIs were automatically segmented using Chan–Vese algorithm. After segmentation, a series of morphology operations were applied to remove noise pixels and to acquire the background pixels surrounding the targets. Then the T/B ratio was calculated as the ratio between the average intensity in the targets and the average intensity in the background. We describe the two main steps in detail below.

2.2.1 Segmentation using Chan–Vese algorithm

The Chan–Vese algorithm uses an active contour model, which initializes the segmentation with a contour and then evolves this contour via a level set method. A level set function defines the edge of the contour. For example, for level set function $\phi(i, j, t)$, where (i, j) is the pixel coordinate, t is the evolution time step, the edge contour is defined to be the zero level set s.t. $\phi(i, j, t) = 0$, while the regions inside and outside the contour are given by $\phi > 0$ and $\phi < 0$. The goal of this algorithm is to evolve the level set function to minimize the energy function as²²

$$F(\phi) = \mu \int_{\Omega} |\nabla H(\phi)| dx + \nu \int_{\Omega} H(\phi) dx + \lambda_1 \int_{\Omega} |I - c_1|^2 H(\phi) dx + \lambda_2 \int_{\Omega} |I - c_2|^2 [1 - H(\phi)] dx, \quad (1)$$

where H is the Heaviside function, ϕ is the level set function, I is the image needed to be segmented, and Ω is the image domain. c_1 and c_2 are the average image intensity in the regions where $\phi > 0$ and $\phi < 0$, respectively [Eqs. (2) and (3)]. μ , ν , λ_1 , and λ_2 are the regulation parameters for the energy function. μ penalizes the length of the edge contour, while ν is the penalty on the total area inside the contour. λ_1 and λ_2 control the variance of pixel intensity inside and outside the contour

$$c_1 = \frac{\int_{\Omega} I(x) H(\phi) dx}{\int_{\Omega} H(\phi) dx}, \quad (2)$$

Fig. 1 T/B ratio calculation procedure for fluorescence image.

$$c_2 = \frac{\int_{\Omega} I(x) [1 - H(\phi)] dx}{\int_{\Omega} [1 - H(\phi)] dx}. \quad (3)$$

To minimize the energy function and to evolve the level set boundary, a semi-implicit gradient descent method is applied.²² The derived numerical and discretized gradients for the level set function [Eq. (4)] minimize the energy function iteratively. The level set function is first initialized with a circular function. At each iteration, ϕ_{n+1} is calculated based on Eq. (4), while c_1 and c_2 are calculated from Eqs. (2) and (3). The iteration will be terminated when the relative change between ϕ_{n+1} and ϕ_n is smaller than a preset threshold or the maximum iteration number is reached, whichever is achieved earlier

$$\begin{aligned} & \frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\Delta t} \\ &= \delta_h(\phi_{i,j}^n) \frac{\mu}{h^2} \left\{ \Delta_-^x \left[\frac{\Delta_+^x \phi_{i,j}^{n+1}}{\sqrt{(\Delta_+^x \phi_{i,j}^n)^2 / h^2 + (\phi_{i,j+1}^n - \phi_{i,j-1}^n)^2 / (2h)^2}} \right] \right. \\ & \quad \left. + \Delta_-^y \left[\frac{\Delta_+^y \phi_{i,j}^{n+1}}{\sqrt{(\phi_{i+1,j}^n - \phi_{i-1,j}^n)^2 / (2h)^2 + (\Delta_+^y \phi_{i,j}^n)^2 / h^2}} \right] \right\} \\ & \quad - \delta_h(\phi_{i,j}^n) \{ \nu + \lambda_1 [I_{i,j} - c_1(\phi^n)]^2 - \lambda_2 [I_{i,j} - c_2(\phi^n)]^2 \}, \end{aligned} \quad (4)$$

where δ_h is the smoothed delta function [Eq. (5)], Δ_- and Δ_+ are the spatial finite difference [Eq. (6)]

$$\delta_h(x) = \frac{1}{\pi} \frac{h}{h^2 + x^2}, \quad (5)$$

$$\begin{aligned} \Delta_+^x \phi_{i,j}^n &= \phi_{i+1,j}^n - \phi_{i,j}^n & \Delta_-^x \phi_{i,j}^n &= \phi_{i,j}^n - \phi_{i-1,j}^n \\ \Delta_+^y \phi_{i,j}^n &= \phi_{i,j+1}^n - \phi_{i,j}^n & \Delta_-^y \phi_{i,j}^n &= \phi_{i,j}^n - \phi_{i,j-1}^n. \end{aligned} \quad (6)$$

2.2.2 Morphology and target/background ratio calculation

The resulting level set function ϕ defines the regions inside and outside the edge contour, which can be transformed into a binary image. Erosion was then applied to smooth the binary image and to remove isolated noise pixels. The default setting is to remove noise smaller than 10 pixels. Then dilation was carried out to generate a 30 pixel-wide background band surrounding the targets. Average intensities of the segmented targets and the neighboring background bands are used for T/B ratio calculation to be able to match results from Joshi et al.¹⁹

Using implementation of the above procedure, we measured the running time of T/B ratio calculation for each frame to evaluate its ability to be used for real-time biopsy guidance.

2.3 Part 2: Automated Computer-Aided Diagnosis for High-Grade Dysplasia and Esophageal Adenocarcinoma

After T/B ratio was calculated for each frame in the fluorescence videos, a histogram-based method was adopted to remove artificial frames (described below) and to determine the T/B ratio of the patients. Then, a cutoff threshold was picked for the best

diagnostic performance to classify EAC/HGD using these recorded fluorescence videos. We introduce each step in more detail in this automated CAD part 2 as follows.

2.3.1 Patients' target/background ratios via histogram-based method

Artifact frames with bright and sometimes colored boundaries are mainly caused by reflection from the optical lens and often result in high T/B ratios. A histogram method was adopted to remove this artifactual effect. T/B ratios of each fluorescence video were first represented as a histogram. Then frames with low or no fluorescence signals ("healthy" frames), which by design cannot be segmented using the Chan–Vese algorithm and have a T/B ratio equal to 1.0, were removed from the histogram. Upper cutoffs derived from the histogram are applied to remove noise and artifactual effects (mentioned above). Since these video frames contain the same regional tissue that is eventually removed by EMR and assigned a single clinical evaluation, statistics (mean and median) of the histogram are applied to acquire the characteristic metrics of all the frames and to compute a single T/B ratio for the patient from all resulting video frames (Fig. 2). This single value represents the T/B ratio for the patient. We tested different combinations of statistics (mean and median) and histogram cutoff (lower 50% to 100%) for comparison. The optimal combination was selected based on the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. For each combination of statistics and histogram cutoff, the T/B ratios for all 50 patients were obtained. By setting different diagnostic cutoff thresholds for the T/B ratio and comparing to the diagnoses based on the most severe histological finding in the patient, we generated ROC curves. The combination with the largest AUC was selected for further experiments and evaluation.

Fig. 2 Histogram-based method to remove artifact frames. A frame with T/B ratio equal to 1.0 was first removed from the histogram, and then a combination of statistics and histogram cutoff is applied for computing a single T/B ratio for the patient. In this example, this single value is calculated as mean T/B ratio of lower 90% of all frames.

2.3.2 Threshold for automatic diagnosis

We used the Youden index²⁴ of the ROC curve to select the optimal cutoff threshold for HGD and EAC classification. Youden index calculates the maximum vertical distance between ROC curve and the diagonal line. Normally, Youden index = sensitivity + specificity – 1.

2.3.3 Classification evaluation

We evaluated our proposed method for automatic EAC/HGD classification using leave-one-out cross-validation (LOOCV). For each round of LOOCV, 49 videos were used as a training set to decide the optimal cutoff using Youden index, which was then tested on the other remaining video. An average cutoff threshold of 50 rounds will be reported, as well as the classification accuracy, the average specificity, and sensitivity over 50 cross-validation tests.

3 Results

3.1 Segmentation of Regions of Interest with Morphology Operation

All 50 fluorescence videos included in this study were able to be automatically segmented to identify the ROIs and to calculate T/B ratios. Each of the video frames was first smoothed with a 5×5 Gaussian filter and then segmented using the Chan–Vese algorithm. The parameters, shown as follows, were taken as baseline, with $\mu = 1$, $\nu = 60$, $\lambda_1 = 2.0$, $\lambda_2 = 2.1$, $dt = 0.5$, and $h = 1$. In the developed user interface, the penalty term ν can be increased to limit area size inside the contour with a slider. Figure 3 shows segmentation (c) and morphology (d and e) results starting from an initial input video frame. Targets smaller than 10 pixels were removed by applying erosion (d). Neighboring contour with a width of 30 pixels is used as background to calculate T/B ratio (e). Red highlighting of the targeted boundary is exhibited in (f).

3.2 Target/Background Ratio Calculation

T/B ratios were computed from labeled ROIs. The following analyses were conducted to quantitatively evaluate the performance of our proposed approach.

3.2.1 Sensitivity analysis

We tested how target and background size affected T/B ratio.

Target size: Due to nonuniform target intensity in certain fluorescence images, changing the constraints on contour area size will change the boundary of the target, thus changing the T/B ratio. Figure 4 shows an example that as we constrained the contour to brighter and smaller targets and maintained the remaining parameters unchanged, the T/B ratio would increase.

Background size: We applied dilation to generate a background band surrounding the tissue. As we include more pixels as background, the average intensity of the background decreases, thus increasing T/B ratio (Fig. 5), which is consistent to all fluorescence videos. This is because the background is nonuniform, with lower intensity further from the target. In this work, we regarded neighboring contour with a width of 30 pixels as background, which is the same standard as Joshi et al.¹⁹

Fig. 3 (a) Raw fluorescence frame, (b) image after 5×5 kernel Gaussian filter, (c) segmentation result with parameter $\mu = 1$, $\nu = 60$, $\lambda_1 = 2.0$, $\lambda_2 = 2.1$, $dt = 0.5$, $h = 1$, (d) image after erosion applied to remove small isolated pixels, (e) image after dilation to generate surrounding background band with a width of 30 pixels (gray band), and (f) highlighted fluorescence target ROI (inside red contour) with T/B ratio greater than the set threshold.

Fig. 4 (a) Fluorescence image collected without processing, (b) T/B ratio 1.48, calculated with parameter $\mu = 1$, $\nu = 60$, $\lambda_1 = 2.0$, $\lambda_2 = 2.1$, $dt = 0.5$, $h = 1$, and $B = 30$ pixels, (c) T/B ratio 1.52, calculated with parameter $\mu = 1$, $\nu = 180$, $\lambda_1 = 2.0$, $\lambda_2 = 2.1$, $dt = 0.5$, $h = 1$, and $B = 30$ pixels. (d) T/B ratio 1.64, calculated with parameter $\mu = 1$, $\nu = 330$, $\lambda_1 = 2.0$, $\lambda_2 = 2.1$, $dt = 0.5$, $h = 1$, and $B = 30$ pixels. Comparing (b) to (d), with other parameters unchanged, more constraints were applied on the area size inside the contour, which resulted in smaller target area and increased T/B ratio.

Fig. 5 (a) Fluorescence image collected without processing, (b) T/B ratio calculated using surrounding background with a width of (b) $B = 20$ pixels, T/B = 1.57, (c) 30 pixels, T/B = 1.64, and (d) 40 pixels, T/B = 1.71. Comparing (b) to (d), with other parameters unchanged, wider surrounding background resulted in higher T/B ratio.

3.2.2 Speed performance

The total processing procedure is implemented in C++ using openCV library. The bulk computation time for calculating T/B ratio for each frame is from the segmentation procedure. For each evolution of the level set function, the computing speed is almost constant. However, to get a reliable segmentation (repeatability >98%), a maximum of eight iterations was used. The average processing time for each frame was 0.5 s. The speed was measured using an Intel Core i7-4980HQ at 2.80-GHz CPU PC with 16-GB memory. Figure 6 shows an example of T/B ratios calculated from continuous video frames (Video 1).

3.3 Histogram-Based Method Performance

AUC is used to select the optimal combination of statistics and histogram cutoff for computing a single T/B ratio for each patient. Based on the T/B ratios for all 50 patients, we calculated the sensitivity and specificity by setting different diagnostic T/B cutoff thresholds, generated ROC curves, and obtained AUCs. AUCs from combinations of statistics and histogram cutoffs (lower 50% to 100%) are shown in Fig. 7. By selecting the maximal AUC of 0.875, the optimal way to compute T/B ratio for each video was determined to be the median of lower 95% frame T/B ratios after removing frames with T/B ratio equivalent to 1.0. Compared to the mean, the median is less susceptible to outliers and extreme values. Differences between AUCs using mean versus median statistics and using histogram cutoff of lower 50% to 100% were not statistically significant, although the sample size of this study limits the power to detect small differences.

3.4 Classification Evaluation

For each patient, T/B ratio was calculated using the best combination of statistics and histogram cutoff selected above, and

T/B Ratio 1.71

Fig. 6 Example of T/B ratios calculated from continuous video frames (Video 1, MP4, 2.14 MB [URL: <http://dx.doi.org/10.1117/1.JMI.4.2.024502.1>]).

Fig. 7 AUC from different statistics and histogram cutoffs to calculate T/B ratio for patients. The best combination to compute T/B ratio for each video is median of lower 95% frame T/B ratios after removing frames with T/B ratio equivalent to 1.0, with AUC 0.875.

classified by the highest grade found on pathology [Fig. 8(a)]. The mean T/B ratio as well as variance for each category is plotted in Fig. 8(b). As we can clearly see, mean T/B ratios for either EAC or HGD were greater than those for BE, GEJ, and LGD. However, the mean T/B ratio for SQ, which is the normal case, is larger than BE and GEJ. This is because there are only 2 SQ cases and one false positive case that is retained in the analysis appeared to bias these results. Different diagnostic cutoff T/B thresholds were set, and sensitivity and specificity curves were generated using T/B ratios for all 50 patients [Fig. 8(c)]. At a target/background ratio of 1.33, which was calculated by the Youden index, there was 94% sensitivity and 75% specificity. At a target/background ratio of 1.45, the specificity was 88%, while the sensitivity was 66%. The ROC curve comparing the T/B ratios for early neoplasia (HGD and EAC) with those for the other classifications (SQ, BE, LGD, and GEJ) achieved an AUC of 0.875. LOOCV was performed to evaluate our proposed method for automatic EAC/HGD classification, which prevents overfitting and combines average of classification error to derive a more accurate estimate of classification performance. The average cutoff threshold calculated from LOOCV is 1.34. The average sensitivity and specificity are 0.92 and 0.75, respectively. The average AUC of the ROC curve is 0.871. With 43/50 accurate classification on one remaining test set, the classification accuracy is 86%. These results are used as the classifier results, with 1.34 as cutoff threshold for highlighting ROIs for biopsy and endoscopic therapy. The sensitivity and specificity for automatic EAC/HGD classification method are 0.92 and 0.75, respectively.

4 Discussion

In this paper, we demonstrated a procedure to calculate T/B ratio in quasireal time using human clinical data that take advantage of a fluorescence peptide and a custom endoscope used *in vivo*. Relative quantification was provided by the amount of molecular probe binding in the target region compared to the surrounding tissue (background), based on the relative fluorescence intensity that is measured from both regions within a video

Fig. 8 (a) T/B ratios for individual patients grouped by the most advanced grade found on pathology. (b) The mean and standard deviation values for T/B ratio in each category. The mean values were 1.44 for SQ ($n = 2$), 1.27 for BE ($n = 3$), 1.26 for GEJ ($n = 4$), 1.35 for LGD ($n = 7$), 1.57 for HGD ($n = 21$), and 1.56 for EAC ($n = 13$). (c) At a target/background ratio of 1.33, sensitivity was 94% and specificity was 75%. At a target/background ratio of 1.45, specificity was 88%, and sensitivity was 66%. These are the results calculated using T/B ratios for all 50 patients. The final results will be calculated as the average from LOOCV (see above). (d) The ROC curve comparing the T/B ratios for early neoplasia (HGD and EAC) with those for the other classifications (SQ, BE, LGD, and GEJ) had an AUC of 0.875.

frame. T/B ratios above a threshold are indicative of progression to EAC from Joshi's previous work.¹⁹ We extend the use of T/B quantification to provide potential real-time highlighting of targeted regions that are above a set threshold. This set threshold is calculated using our automatic CAD algorithm (part 2) for classifying cancer (EAC) and its precursor (HGD). LOOCV is performed to evaluate this method and to prevent overfitting. The average T/B threshold from LOOCV, 1.34, is used as a cut-off threshold for highlighting ROIs for biopsy and endoscopic therapy. A user interface is designed to highlight the targets by enhancing the color of the targets or simply encircling the targets with a red boundary and the latter was chosen by the coauthoring gastroenterologists (Fig. 9). A toggle switch is being designed to change the cutoff threshold depending on the clinical scenario, routine surveillance (biopsies), or therapy (EMR). The design iterations are continuing for this clinical interface. The initial processing speed is 2 fps, while the ideal processing speed is to match the frame rate of the endoscopy, which is above 5 fps. The processing speed can be accelerated using graphics processing unit parallel processing. This processing speed can be further accelerated by using field-programmable gate array (FPGA) architecture, which is available in our developing scanning fiber endoscope (SFE see below).

The most time-consuming part of the processing is the segmentation step. Instead of processing the iterative evolution of the level set function with only one processor, FPGA can pipeline each iteration with multiple processors. In addition, using FPGA parallel processing makes the acceleration possible. This method demonstration is designed to help an experienced endoscopist to locate and track suspected targets measured to be the highest risk for cancer for subsequent biopsy or surgical removal.

We developed an automatic CAD algorithm for classifying esophageal neoplasia (part 2), using the same recorded endoscopic video frames from Joshi et al.,¹⁹ which captured fluorescence from targeted tissues being removed by EMR or biopsy, and confirmed by histopathology. This unique data set consists of a series of video frames from the same region of targeted tissue that is eventually removed and assigned a single clinical evaluation. By analyzing the series of video frames for each case, an overall diagnostic T/B value can be assigned for each patient. In calculating this overall diagnostic T/B value, repeated measures of the T/B ratio are made for each video frame. This allows a sensitivity analysis under clinical conditions of a range of variables, which test the robustness of our image processing algorithms. Analysis on surrounding background has been

Fig. 9 Two methods to highlight the targets: (a) enhancing the color of the targets and (b) encircling the targets.

applied to test the sensitivity. When including more pixels as background, T/B ratio increases. To mitigate the effect of background size on T/B ratio, we regarded neighboring contour with a width of 30 pixels to match Joshi et al.'s¹⁹ postprocessing study results, which are compared below (Table 1) to our automated CAD approach. The future clinical utility of this method is expected to aid a less experienced endoscopist in the referral decision making process for suspicion of EAC.

We can see in Table 1 that our overall mean T/B ratios for each category are smaller than those for Joshi's, except for the SQ case. The variables or clinical issues that make this difference in T/B ratio between our automated method and Joshi's method using the same video data set are assumed to be (1) different segmentation algorithms to label fluorescent targets, (2) user selection of video frames by Joshi et al., and (3) inconsistent distance between the detector and the imaged mucosa.¹⁹ The Otsu's segmentation algorithm, which Joshi applied, is limited by excluding small targets with lower intensity from the labeled target region, since this algorithm applies global threshold to segment the image. Instead, we applied Chan–Vese algorithm, which is more robust to noisy, low-contrast images. Our segmentation algorithm includes more small targets with lower intensity, thus decreasing the T/B ratio. Moreover, Joshi selected video frames with high quality for postprocessing, while we calculated each frame in the videos. So there is no interobserver variation in selecting the frames in our method, there will be some systematic differences in original content when comparing to Joshi et al. Finally, to compensate for uncontrolled or

inconsistent distance between the detector and the imaged mucosa, Joshi et al.¹⁹ generated a ratiometric image (fluorescence image/reflectance image) before calculating the T/B ratio. However, since the fluorescence image and the reflectance image are not captured concurrently and there is registration artifact when performing imaging procedure in time series, it is impossible to register all frames in the fluorescence video with the reflectance video. The ratio image suppresses the noise and increases the contrast of the image thus increasing the T/B ratio. However, our mean T/B ratio for SQ is 19% greater than Joshi's results. The main reason is assumed to be the small sample size of the SQ cases and one false positive case among the two SQ cases biases the results. At a T/B ratio of 1.34, the average sensitivity and specificity from LOOCV using our method are 92% and 75%, respectively, compared to Joshi's of 76% and 94% at a T/B ratio at 1.49, respectively (described in Table 2).

Even though we have different cutoff threshold, sensitivity, and specificity compared to Joshi's, we returned a similar AUC for the ROC curve, 0.871 compared to Joshi's 0.884,¹⁹ which indicates both methods can be applied to calculate T/B ratio for patients. The difference in sensitivity and specificity lies in different calculated T/B ratios, different selected cutoff threshold, and systematic difference. In this paper, the cutoff threshold for the ROC curve was calculated using the Youden index. However, in practice, the cutoff threshold should be incorporated with the clinical scenario for correct and false diagnosis. For example, when taking routine surveillance and

Table 1 Comparison between mean T/B ratio for individual patients grouped by histopathology classification.

	SQ (<i>n</i> = 2)	BE (<i>n</i> = 3)	GEJ (<i>n</i> = 4)	LGD (<i>n</i> = 7)	HGD (<i>n</i> = 21)	EAC (<i>n</i> = 13)
Joshi's results	1.21	1.39	1.27	1.39	1.73	1.65
Our results	1.44	1.27	1.26	1.35	1.57	1.56
Difference (%)	19.0	−8.6	−0.8	−2.9	−9.2	−5.5

Note: SQ, squamous; BE, Barrett's esophagus; GEJ, gastroesophageal junction; LGD, low-grade dysplasia; HGD, high-grade dysplasia; and EAC, esophageal adenocarcinoma.

Table 2 Comparison between results from our automated CAD method and Joshi's postprocessing CAD method.

	T/B cutoff threshold	Sensitivity (%)	Specificity (%)
Joshi's results	1.49	76	94
Our results	1.34	92	75
PIVI recommendations	—	90	80

guiding biopsies, the cutoff threshold should obtain a high sensitivity to rule out the possibility of false negative. When performing endoscopic resection, a procedure with greater risks to the patient than biopsies, the cutoff threshold should generate a higher specificity to decrease the possibility of false positive. Overall, we have reasonable preliminary results for a future EAC surveillance program, compared to Joshi's. As stated in the *Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI)* report, the new imaging technology should have a recommended sensitivity of 90% or greater and maintain a specificity at least 80% for detecting HGD and EAC during the endoscopic surveillance of patients with BE.²⁵ Our automated CAD for EAC and HGD met the criterion for sensitivity at 92% but failed to meet the criterion for specificity at 75% in this preliminary study with 50 BE subjects.

One of the uncontrolled variables in our approach lies in the unfixed segmented boundary. Therefore, a change in the penalty on the area inside the contour will result in a change in segmented target size, thus changing the T/B ratio in some frames as shown in Sec. 3.2.1. Clinical tests should be conducted to evaluate the expected range of target sizes of HGD lesions for EAC, which may be different for subpopulations and of course other cancers. Another uncontrolled variable is that there is no distance compensation in our approach, which can assist in removing artifacts and noise in the image and enhancing image contrast. Moreover, we only applied a single fluorescence peptide probe to detect one biomarker. Last but not least, we calculate the average T/B ratio of multiple targets to represent the T/B ratio of the frame. If the intensity for the targets is not uniform, then the average T/B ratio underestimates the brighter targets. However, this issue can be resolved by the endoscopist

adjusting the position of the scope closer to the brighter targeted tissue, as our method has the potential to recalculate T/B ratio in real time under these dynamic conditions.

To overcome some of these limitations, we are developing the SFE for first-in-human multimodal endoscopic procedures using fluorescence peptides from Dr. TD Wang's group at the University of Michigan. Since the SFE has an ultrathin and flexible profile, there can be a range of different manifestations to reach the esophagus, such as a mini-endoscope within the working channel of a conventional endoscope or as a transnasal design or tethered-capsule design.²⁶ Compared to the much larger caliber Olympus system, our SFE can acquire fluorescence and reflectance videos concurrently, which can account for the inconsistent distance between the detector and the imaged mucosa.²⁷ Furthermore, other systematic errors can be reduced with multimodal SFE imaging, such as real-time autofluorescence background mitigation algorithm and elimination of specular reflection artifacts, thus increasing the accuracy and scale of the T/B value by over an order of magnitude.^{28,29} These improvements are expected to increase the specificity to reach the desired minimum of 80% (PIVI recommendations, Table 2) while maintaining automation of the analysis. Moreover, the SFE has multiple channels in a wide range of detection wavelengths (visible to infrared), which makes multiple molecular targets detection possible. Additionally, the FPGA architecture in the SFE makes real-time guiding biopsy using multiple fluorescence channels possible. Multiple molecular probe imaging of early cancer is expected to increase sensitivity without sacrificing specificity due to the high-molecular specificity of the peptide fluorescence probes.³⁰ Comparison between Joshi's method w/Olympus, our method w/Olympus, and future work in esophagus w/SFE is described in Table 3.

In our future work using either multimodal endoscopy system, we aim to calculate the T/B ratio in real time. The combination of SFE and automatic, real-time fluorescence image analysis is estimated to greatly improve the efficiency and safety of guiding biopsy and surgical removal of neoplasia in the esophagus and smaller ductal regions.

Disclosures

Author E.J.S. participates in royalty sharing with his employer, the University of Washington, which has ownership of patents that may gain or lose financially through this publication.

Table 3 Joshi's method w/Olympus, our method w/Olympus, and future work in esophagus w/SFE.

	Olympus + Joshi's method	Olympus + our method	SFE + our method
CAD with T/B ratio	Yes	Yes and automatic	Yes and automatic
Segmentation algorithm for T/B calculation	Otsu's algorithm: single threshold	Chan-Vese: active contour, robust to patchy, noisy images	Chan-Vese: active contour, robust to patchy, noisy images
Video frame selection	Manual	All	All
Fluorescence and reflectance capture	Time series	Time series	Concurrent
Distance compensation	Yes	No	Yes
Multiple fluorescence channels	No	No	Yes
Real-time capability	Postprocessing	2 fps	5 to 30 fps

Acknowledgments

This study was supported in part by the National Institutes of Health U54 CA163059, which is part of the NCI Barrett's Esophagus Translational Research Network, and NIH R01 CA200007 (PI's are T.D.W. and E.J.S.). Helpful discussions to the authors were provided by Dr. Bishnu Joshi.

References

1. N. Howlader et al., *SEER Cancer Statistics Review, 1975–2013*, National Cancer Institute, Rockville, Maryland (2006).
2. F. Hvid-Jensen et al., "Incidence of adenocarcinoma among patients with Barrett's esophagus," *N. Engl. J. Med.* **365**(15), 1375–1383 (2011).
3. S. J. Spechler and R. K. Goyal, "The columnar-lined esophagus, intestinal metaplasia, and Norman Barrett," *Gastroenterology* **110**(2), 614–621 (1996).
4. S. Bhat et al., "Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study," *J. Natl. Cancer Inst.* **103**(13), 1049–1057 (2011).
5. N. J. Shaheen et al., "ACG clinical guideline: diagnosis and management of Barrett's esophagus," *Am. J. Gastroenterol.* **111**(1), 30–50 (2016).
6. J. A. Abrams et al., "Adherence to biopsy guidelines for Barrett's esophagus surveillance in the community setting in the United States," *Clin. Gastroenterol. Hepatol.* **7**(7), 736–742 (2009).
7. D. F. Boerwinkel et al., "Effects of autofluorescence imaging on detection and treatment of early neoplasia in patients with Barrett's esophagus," *Clin. Gastroenterol. Hepatol.* **12**(5), 774–781 (2014).
8. H. C. Wolfsen et al., "Prospective, controlled tandem endoscopy study of narrow band imaging for dysplasia detection in Barrett's esophagus," *Gastroenterology* **135**(1), 24–31 (2008).
9. M. J. Gora et al., "Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure," *Nat. Med.* **19**(2), 238–240 (2013).
10. D. F. Boerwinkel et al., "Third-generation autofluorescence endoscopy for the detection of early neoplasia in Barrett's esophagus: a pilot study," *Dis. Esophagus* **27**(3), 276–284 (2014).
11. M. Muto et al., "Narrow band imaging: a new diagnostic approach to visualize angiogenesis in superficial neoplasia," *Clin. Gastroenterol. Hepatol.* **3**(7), S16–S20 (2005).
12. W. L. Curvers et al., "Endoscopic tri-modal imaging for detection of early neoplasia in Barrett's oesophagus: a multi-centre feasibility study using high-resolution endoscopy, autofluorescence imaging and narrow band imaging incorporated in one endoscopy system," *Gut* **57**(2), 167–172 (2008).
13. N. G. Terry et al., "Detection of dysplasia in Barrett's esophagus with in vivo depth-resolved nuclear morphology measurements," *Gastroenterology* **140**(1), 42–50 (2011).
14. H. C. Wolfsen et al., "Safety and feasibility of volumetric laser endomicroscopy in patients with Barrett's esophagus (with videos)," *Gastrointest. Endoscopy* **82**(4), 631–640 (2015).
15. M. B. Sturm et al., "Targeted imaging of esophageal neoplasia with a fluorescently labeled peptide: first-in-human results," *Sci. Transl. Med.* **5**(184), 184ra61 (2013).
16. P. L. Hsiung et al., "Detection of colonic dysplasia in vivo using a targeted heptapeptide and confocal microendoscopy," *Nat. Med.* **14**(4), 454–458 (2008).
17. C. L. Leggett et al., "Comparative diagnostic performance of volumetric laser endomicroscopy and confocal laser endomicroscopy in the detection of dysplasia associated with Barrett's esophagus," *Gastrointest. Endoscopy* **83**(5), 880–888 (2016).
18. E. L. Bird-Lieberman et al., "Molecular imaging using fluorescent lectins permits rapid endoscopic identification of dysplasia in Barrett's esophagus," *Nat. Med.* **18**(2), 315–321 (2012).
19. B. P. Joshi et al., "Multimodal endoscope can quantify wide-field fluorescence detection of Barrett's neoplasia," *Endoscopy* **48**(2), A1–A13 (2016).
20. X. Qi et al., "Computer-aided diagnosis of dysplasia in Barrett's esophagus using endoscopic optical coherence tomography," *J. Biomed. Opt.* **11**(4), 044010 (2006).
21. P. K. Chan, S. H. Cheng, and T. C. Poon, "Automated segmentation in confocal images using a density clustering method," *J. Electron. Imaging* **16**(4), 043003 (2007).
22. T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
23. Y. Jiang et al., "Real-time fluorescence target/background (T/B) ratio calculation in multimodal endoscopy for detecting GI tract cancer," *Proc. SPIE* **10057**, 100570D (2017).
24. K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation," *Caspian J. Intern. Med.* **4**(2), 627–635 (2013).
25. P. Sharma et al., "The American Society for Gastrointestinal Endoscopy PIVI (preservation and incorporation of valuable endoscopic innovations) on imaging in Barrett's esophagus," *Gastrointest. Endoscopy* **76**(2), 252–254 (2012).
26. C. M. Lee et al., "Scanning fiber endoscopy with highly flexible, 1-mm catheterscopes for wide-field, full-color imaging," *J. Biophotonics* **3**(5–6), 385–407 (2010).
27. C. Yang et al., "Color-matched and fluorescence-labeled esophagus phantom and its applications," *J. Biomed. Opt.* **18**(2), 026020 (2013).
28. C. Yang et al., "Mitigating fluorescence spectral overlap in wide-field endoscopic imaging," *J. Biomed. Opt.* **18**(8), 086012 (2013).
29. C. Yang et al., "Target-to-background enhancement in multispectral endoscopy with background autofluorescence mitigation for quantitative molecular imaging," *J. Biomed. Opt.* **19**(7), 076014 (2014).
30. B. P. Joshi et al., "Multispectral endoscopic imaging of colorectal dysplasia in vivo," *Gastroenterology* **143**(6), 1435–1437 (2012).

Biographies for the authors are not available.