1. Introduction

Bleeding is one of the most common lesions about gastrointestinal (GI) disease [1]. The conventional detection method requires physicians to scan real-time images through an endoscope-attached wire, which is clinically invasive. Capsule endoscope camera, on the distal side of which a miniaturized sensor is mounted, can sweep past patients’ gut wall. Such noninvasive endoscopy, called wireless capsule endoscopy (WCE), can capture a video of nearly 57,000 frames, where the bleeding locations are recorded. Detecting and interpreting the lesions from numerous frames need physicians to be concentrated [2]. To ease the task of subjective diagnosis, computer-assisted bleeding localization systems are proposed [3].

The state-of-the-art bleeding localization systems depend mainly on color features. For instance, authors in [4] extracted mean and variance in the HSI color space: hue, saturation, and intensity. In [5], high-order statistical moments including kurtosis were added to the RGB color space. Similarly, statistical features based on intensity histogram [6] and local binary pattern are computed in the RGB and HSV color space. Unlike color features, edge was discarded since it increased false positive (FP) rate [7]. More importantly, these systems generally compose these statistical features from different color spaces into a vector, followed by a supervised pattern classification tool such as artificial neural network or support vector machine. As a result, the task of localization is degraded since classification fails to predict accurate positions of a lesion.

This paper proposes a saliency detection method that is devoted to localizing the bleeding regions in WCE images. Visual saliency models that extract the most interesting part from an image have been proposed in the nature image recognitions [8–11]. However, it is undesirable to apply the models directly to medical image analysis. Firstly, most existing saliency models assume that the visual fixation tends to stay within the center of images [12], while, in medical images, the fixation mainly refers to lesions which appear anywhere in the field of view. Secondly, salient regions in the natural images tend to be in the foreground where edge and color information is quite distinguishable; however, lesions in the WCE images vary according to the specific diseases. For instance, the bleeding lesions are small and dark red, while the colon lesions appear large and highly laminated. Thirdly, in the bleeding localization system, the salient region must exclude not only normal tissue, but also gastric fluid and undigested residue. The latter two are quite different from the...
normal, which can prevent the system from recognizing FP samples.

To detect the bleeding regions with both high sensitivity and specificity, the traditional methods take full advantage of color features and ignore other features. However, when diagnosing the bleeding area, physicians use multiple features, such as edge, intensity, and color. In order to simulate the diagnosis process, a visual saliency-based WCE bleeding detection system is proposed, which includes three kinds of saliency maps, color, intensity, and edge. The main contribution of this paper is that intensity and edge saliency maps are introduced to better imitate the diagnosis of physician. Salient intensity is extracted from the intensity histograms which have rotation invariance while edges are selected based on the phase congruency [9] of Log-Gabor filter bank which can solve the problem of nonuniform illumination of the WCE. Experimental results demonstrate that our algorithm has a very good performance for detecting the bleeding area. The rest of this paper is organized as follows. Materials and methods are presented in Section 2; Section 3 exhibits the experimental results; finally, some conclusion is drawn in Section 4.

2. Materials

A total of 200 WCE frames, with the pixel resolution of 480-by-480, are used with a warrant from Ankon Incorporation, Wuhan, China. Two physicians, with three-year clinical experience, drew bleeding masks independently; the two masks of each image are fused together by pixel-to-pixel AND operation. Randomly selected six bleeding images and masks are presented in Figure 1.

3. Saliency Map Extraction

Clinicians discriminate the bleeding lesions from a WCE image mainly based on some salient information. Here we classify the information into three categories: edge information, intensity information, and color information, each of which is quantified using a saliency map.

3.1. Edge Saliency Map. When clinicians calibrate the bleeding location, the clear contour of bleeding area is an important reference condition. Therefore, the outlines of bleeding area will occupy a large proportion in the clinician’s attention mechanism to distinguish the bleeding area from the normal. In order to simulate the diagnosis of clinician, we propose the edge saliency map.

Generally, edges can be obtained from spatial edge detectors such as Roberts, Laplace, or Canny, but all these detectors are vulnerable to the image noise. Edge detection based on PC has the brightness invariance properties, which can compensate for the instability luminance of color channels [13]. The solution of extracting noise-insensitive edge is to convert an image into the spectrum domain and to calculate the phase congruency (PC [9]). Still, there are many frequency transformation tools like Fast Fourier Transformation (FFT), discrete cosine transformation (DCT), and Gabor filters. However, the FFT and DCT fail to be analyzed in the multiresolution of scales and orientations, which make the PC too coarse. Though Gabor filters can be used in the different scales and orientation, they tend to be bias frequency components towards the lowest band. Hence, the PCs calculated from Gabor frequency bands are not isolated. Here we choose the Log-Gabor wavelet transform because it is competent to extract the PC at the isolated center frequencies with symmetry attenuation responses [14]. The scales and orientations parameters are empirically set to 5 and 6, respectively. According to the spectrum analysis, an image can be decomposed into a combination of amplitude spectrum and phase spectrum. Accordingly, the Log-Gabor filter is constructed by multiplying the frequency response of the two components together in the polar coordinates system as follows:

\[
\text{LG}(\rho, \theta) = \exp\left\{-\frac{\log(\rho/f_0)}{2\sigma_\rho^2}\right\} \cdot \exp\left\{-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2}\right\},
\]

where \((\rho, \theta)\) represents the polar coordinates and \(f_0\) is the center frequency of the filter and it is related to our current scale \(n\) by \(f_0 = \text{minWave} \times \text{mult}^n\) in which \text{minWave} is the wavelength of smallest scale filter and mult is the scaling factor between the successive filters. \(\theta_0\) is the orientation angle of the filter and \(\sigma_\rho\) and \(\sigma_\theta\) determine the scale bandwidth and the angular bandwidth, respectively. In our experiments, the parameters for the Log-Gabor filters are set as follows: \text{minWave} = 6.0, \text{mult} = 2, \sigma_\rho = 0.75, \sigma_\theta = 0.6.
At each orientation, a PC map in the \( n \)th scales is fused by
\[
\text{PC}(x, y) = \sqrt{\left(\sum_{n} I(x, y) \ast R_n\right)^2 + \left(\sum_{n} I(x, y) \ast M_n\right)^2} + \varepsilon \sum_{n} \sqrt{\left(I(x, y) \ast R_n\right)^2 + \left(I(x, y) \ast M_n\right)^2},
\]
where \( R_n \) and \( M_n \) denote the even and odd spectrums at the \( n \)th scale. The constant \( \varepsilon \) stabilizes the denominator. The \( \varepsilon \) value of 0.01 is used for all the results presented in this paper.

Now five interscale PC maps at each orientation are fused together. The edge saliency map is obtained by maximizing the singular value-based moment of PC maps at six orientations:
\[
C_x(x, y) = \sum_{d} \left[ \text{PC}(x, y) \times \sin(\theta) \right]^2, \\
C_y(x, y) = \sum_{d} \left[ \text{PC}(x, y) \times \cos(\theta) \right]^2, \\
C_{xy}(x, y) = \sum_{d} \left[ \text{PC}(x, y) \cos(\theta) \times \text{PC}(x, y) \sin(\theta) \right], \\
S_1(x, y) = C_x(x, y) + C_y(x, y) + \sqrt{C_{xy}(x, y)} \left(C_x(x, y) - C_y(x, y)\right)^2,
\]
where \( \theta \) denotes the orientation angle. \( S_1(x) \) is the edge saliency map. Figure 2 provides the flowchart of extracting the PC-based edge saliency map, and Figure 4(b) visualizes the final high-resolution edge saliency maps.

### 3.2. Intensity Saliency Map

The PC-based edge saliency map is sensitive to bleeding contours. On the other hand, the luminance intensities inside the contours are also prone to saliency. The difference in luminance is an indispensable condition for clinicians to distinguish between normal tissue and diseased tissue. Based on this principle, the gray-scale image histogram is exploited to compute an intensity saliency map. The histogram shown in Figure 3 is built by counting the gray-scale intensities within the bins. The counts of each bin are normalized to be a probabilistic value \( (p_k) \) such that
\[
\sum_{k=0}^{N} p_k = 1,
\]
where \( k \) denotes the \( k \)th bin. There are totally \( N \) bins, and, in this paper, \( N = 256 \), which means that one gray-scale corresponds to one bin.

\[
S_2(x, y) = \begin{cases} 
0 & \text{if } P_{(x,y)} > \bar{P} \\
0.25 & \text{if } \lambda \bar{P} < P_{(x,y)} \leq \bar{P} \\
0.5 & \text{if } \lambda^2 \bar{P} < P_{(x,y)} \leq \lambda \bar{P} \\
0.75 & \text{if } \lambda^3 \bar{P} < P_{(x,y)} \leq \lambda^2 \bar{P} \\
1 & \text{if } \lambda^4 \bar{P} \leq P_{(x,y)} 
\end{cases}
\]
where \( \bar{P} \) and \( \lambda \) are set to 1/72 and 0.6, empirically; \( (x, y) \) denotes the pixel position in the gray-scale image \( I \). The function values constitute the intensity saliency map, which is visualized in Figure 4(c).

### 3.3. Color Saliency Map

Besides the above two features, the color feature which is the most widely used in the bleeding area detection cannot be ignored. The obvious color information of the bleeding area is the most intuitive reference for the clinicians, so the state-of-the-art methods utilize the color feature to distinguish the bleeding from other regions [5, 15–17]. In perspective of the visual saliency, color is another visual stimulus that is isotropic to the contour and luminance. Hence color saliency map is also computed and fused together with the above two maps. As most bleeding lesions appear to be red, we calculate the proportion of the red of the color saliency map,
\[
S_3(x, y) = \frac{r(x, y)}{r(x, y) + g(x, y) + b(x, y) + \varepsilon},
\]
where \( r(x, y) \), \( g(x, y) \), and \( b(x, y) \) denote the red, green, and blue components of the pixel \( (x, y) \) in the RGB color space, respectively.
where \([r, g, b]\) is a color vector at the pixel position of \((x, y)\) and \(\varepsilon\) is a small constant for the purpose of stabilization.

The aforementioned three saliency maps localize the bleeding ROIs in WCE images from different aspects. More importantly, the quantities in all the three saliency maps are between 0 and 1; thus we can derive the fused saliency map by fusing them together by the following formula, and the three saliency maps are pooled by

\[
S(x, y) = \frac{1}{3} \sum_{i=1}^{3} S_i(x, y).
\]

As an illustration, Figure 4 presents the three test WCE images in Figure 4(a). For each image, three isolated saliency maps and the contrastive fused saliency maps are arranged from Figures 4(b)–4(h). Figure 4(i) localizes the suspected bleeding ROIs by thresholding the fused saliency map.

\section{4. Experimental Results and Discussion}

To evaluate our saliency-based bleeding detection system, we compared the saliency maps with the golden-standard masks that are marked by two physicians. As each saliency map implies the probability of bleeding positions, a receiver operator characteristic (ROC) curve can illustrate both sensitivity and specificity of a map. Both criteria are defined as follows:

\[
\text{Sensitivity} = \frac{TP}{TP + FN}, \quad \text{Specificity} = \frac{TN}{FP + TN},
\]

where TP denotes true positive rates and FP false positive rates; TN and FN denote true negative and false negative, respectively. TP means that a bleeding pixel is correctly classified by a saliency map while FP means that a nonbleeding pixel is incorrectly regarded as bleeding.

\subsection{4.1. Quantitative Analysis}

Given that a threshold increased from 0 to 1, a series of sensitivity and specificity values are obtained. Relying on these values, a ROC curve can be plotted. As an illustration, Figure 5 plots the ROC curves of fused saliency maps with regard to the six images in Figure 1. From Figure 5, it can be observed that the fused saliency map shows unbalanced performance on the six representative bleeding images. Area under curves (AUC) is 0.994, 0.989, 0.986, 0.926, 0.895, and 0.869, respectively. Note that the fused saliency map tends to be more sensitive to tiny and sharp bleeding region (Figure 1(f)) than opaque one (Figure 1(a)). This may result from the fusion strategy which weights three isotropic maps equivalently.

To further evaluate the performance of the proposed saliency map, the three kinds of saliency maps, saliency maps of different combinations, and the peer algorithm [17] are listed in Table 1, where accuracy, sensitivity, and specificity values for different saliency maps are presented. The accuracy criterion is computed as

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.
\]
where TP, TN, FP, and FN are obtained based on the threshold that is the mean value.

It can be observed that edge and intensity saliency maps are helpful to enhance bleeding detection accuracy. The edge saliency map has high degree of sensitivity to improve the sensitivity of the fused saliency map and the higher specificity of the fused saliency map benefits from the intensity saliency map; it is clear that the proposed saliency map is better than the saliency map in [17], which uses single color features. In order to discuss the robustness of the proposed saliency map, the histogram is drawn in Figure 6. From Figure 6(a), it can be seen that the sensitivity of edge saliency map is better than the other saliency maps; the specificity in Figure 6(b) illustrates that effect of the fused saliency map is very ideal and the values are very close to 1, while the main contribution of specificity of the fused saliency map is from the intensity saliency map; Figure 6(c) shows that the accuracy of the fused saliency map is the best, which meets the expectations of our design.

In consideration of the impact of the color space [18–20], the fused saliency map is applied to the different color spaces in Table 2. From Table 2, the fused saliency map in the RGB color space has better performance than the HSI and the HSV.

In addition, the proposed saliency map provides the probability of bleeding, which is suitable to be used as...
Based on the saliency map, unimportant superpixels can be discarded.

4.2. Qualitative Analysis. In order to intuitively see the effect of the saliency map, we delineate the ROI of bleeding point. As shown in Figure 7, there is the binary image next to each saliency map. The saliency map is binarized according to single thresholding. The threshold of a saliency map is determined by the mean value of all quantities in the saliency map.

It can be seen that color binary mask is oversized while the shape binary mask is undersized. The edge binary mask tends to reveal all potential bleeding regions. Instead of applying logic AND or OR operations to fusing the three isolated binary masks, the proposed fusion strategy performs probabilistic calculation on isolated saliency maps, followed by binarizing the fused saliency map, which increases the fidelity of TP regions. It is obvious to see in Figure 7 that the fused binary map reveals exactly where the bleeding point is.

4.3. Image Noise Influence. WCE images are sometimes vulnerable to the noise contaminations such as Gaussian noise or Salt and Pepper noise. To evaluate the performance of the proposed saliency map, we add three levels of Gaussian noise and Salt and Pepper noise in the WCE image in Figure 8.

From Figure 9 and Table 3, we can find that the accuracy of the proposed saliency map decreases as the level of noise increases, but the accuracy is still in expected range.

4.4. Run-Time Evaluation. The computation time of the fused saliency map and the contrast algorithm [17] is listed in Table 4. The algorithms are run on the MATLAB R2015b, with Intel Core i5 CPU at 2.4 GHz. Compared with [17], the proposed algorithm reduces the algorithm complexity.
of color saliency map and extracts three saliency maps in parallel. Table 4 shows that both algorithms are very fast and the proposed one has a minor advantage.

5. Conclusion

In this paper, a novel saliency map is proposed for bleeding localization on the WCE diagnosis. Unlike existing methods that use color features as the dominant criterion, we combine edge, intensity, and color information in visual saliency scheme. The results demonstrated that both the edge saliency map based on phase congruency and the intensity map based on luminance histogram dominate the saliency detection performance. Fused saliency map can detect the bleeding in WCE images with the average accuracy of 98.97%. Future work involves segmenting bleeding region with precise contours at a high specificity.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Grant no. J20141237). The authors would like to thank Hao Zhang, the project manager of Ankon Incorporation, Wuhan, China.

References


Table 4: Computation time comparison.

<table>
<thead>
<tr>
<th></th>
<th>The fused saliency map [17]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time/ per (s)</td>
<td>0.0748</td>
</tr>
<tr>
<td></td>
<td>0.0855</td>
</tr>
</tbody>
</table>


