2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

# A Novel Method for Capsule Endoscopy Video Automatic Segmentation

Ran Zhou, Baopu Li, *Member, IEEE*, Hongmei Zhu, *Student Member*, *IEEE*,

Max Q.-H Meng, *Fellow, IEEE*

***Abstract-*** **Wireless capsule endoscopy (WCE) is a recently developed revolutionary medical technology which records the video of human's digestive tract noninvasively. However, reviewing a WCE video is a tired and time-consuming task for clinicians. Thus, WCE video automatic segmentation methods are emerging to reduce the review time for clinicians. In our previous work, a two-level WCE video segmentation approach has been proposed, which provides a novel approach to localize the boundaries more exactly and efficiently. However, it has an unsatisfactory performance in the small intestine/large intestine boundary detection. In this paper, we propose new features and an improved classifier to improve the previous two-level segmentation algorithm. In the rough level, color feature is utilized to draw a dissimilarity curve and an approximate boundary has been obtained. At the same time, training data for fine level can be directly labeled and collected between the two approximate boundaries of organs to overcome the difficulty of training data acquisition. In the fine level, a novel color uniform local binary pattern (CULBP) algorithm is proposed, which includes two kinds of patterns, color norm patterns and color angle patterns. The CULBP feature is more robust to variation of illumination and more discriminative for classification. Moreover, in order to elevate the performance of SVM classifier we proposed the Ada-SVM classifier which using RBFSVMs as component of Adaboost classifier. At last, an analysis of classification results of the Ada-SVM classifier is carried out to segment the WCE video into several meaningful parts, stomach, small intestine and large intestine. The experiments demonstrate a promising performance of the proposed method. The average precision and recall are as high as 91.37% and 88.50% in stomach/small intestine classification, 90.35% and 97.28% in small intestine/ large intestine classification.**

***Index terms-*** **Capsule endoscopy, WCE video segmentation, Color uniform local binary pattern, Ada-SVM.**

## I. Introduction

Wireless capsule endoscopy is a revolutionary medical technology which can directly view the human's small intestine noninvasively [1]. The WCE looks like a pill which contains a tiny camera and a wireless communication system. Videos of the digestive tract are captured as the capsule goes through the tract by normal peristalsis which provides a direct visual and non-invasive procedure for clinicians to exam diseases existed in the digestive tract. So WCE has been widely used in hospitals [2]. And examining diseased in the small intestine is the most important task, since which cannot be viewed by using traditional medical technology. However, a WCE video lasts over 8 hours which includes nearly 60,000 frames. Reviewing the WCE videos is a boring and tired task, and it costs even an experienced clinician about two hours to review and analyze a video data [3]. Segmentation of the whole video may help the clinician ensure relevant organ section, estimate the transition time of a WCE and reduce the review time [6].

Many approaches for WCE video automatic segmentation spring up to ease clinicians' burden of reviewing and analyzing videos. Berens et al. [4] utilized hue saturation chromaticity histograms to automatically discriminate stomach, intestine and colon tissue in order to significantly reduce the video assessment time. Lee at al. [5] presented an algorithm to detect event boundaries in WCE videos based on the energy of contractions in frequency domain. In [6] MPEG-7 descriptors and SVM classifier was proposed for automated topographic segmentation in endoscopic capsule exams by Coimbra and Campos. But in the global model fitting step, it needs to estimate and judge all frames in a WCE video, which is a time-consuming procedure. Mackiewicz et al. [7] extracted color and texture features of frames and segmented the video into several meaningful parts using support vector or multivariate Gaussian classifiers. However, acquiring various training data is also a hard work, and the performance of SVM classifier is impacted by parameters. Li and Meng proposed a new method to detect boundaries of WCE videos without using any classifiers[8], in which color and texture features were applied to draw a dissimilarity curve between frames. Furthermore, they investigated the possibility of applying motion analysis approaches for WCE video segmentation in [9]. However, false boundaries detection, such as unusual events of capsule, may impact the segmentation results. In a very recent paper [10], Shen et al. proposed an unsupervised learning approach for WCE video segmentation, in which Scale Invariant Feature Transform (SIFT) was utilized to extract local image features， but SIFT feature extraction is time consuming.

In our previous work [11], a two-level WCE video segmentation algorithm is proposed, which utilizes a pre-process approach and provides a novel method to segment the WCE videos more exactly and efficiently. Moreover, it solves the difficulty of acquiring and labeling diverse training data. However, the performance of segmentation in small intestine/large intestine is not satisfied. Based upon our previous work, this paper attempts to investigate new features and more effective classifier in order to improve the performance in the fine level segmentation, where a novel

Ran Zhou, Baopu Li, Hongmei Zhu are with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China(corresponding author e-mail: bp.li@ee.cuhk.edu.hk)

Max Q.-H Meng and Baopu Li are with Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, China

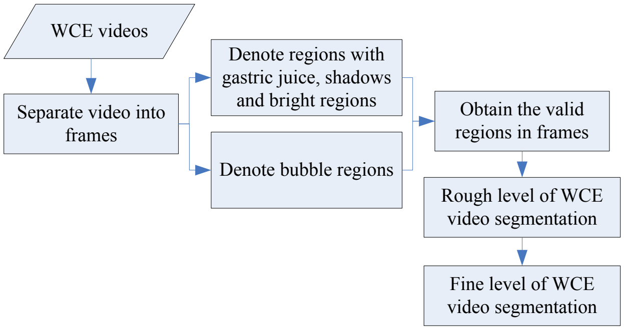


Fig.1 Flowchart of our previous work

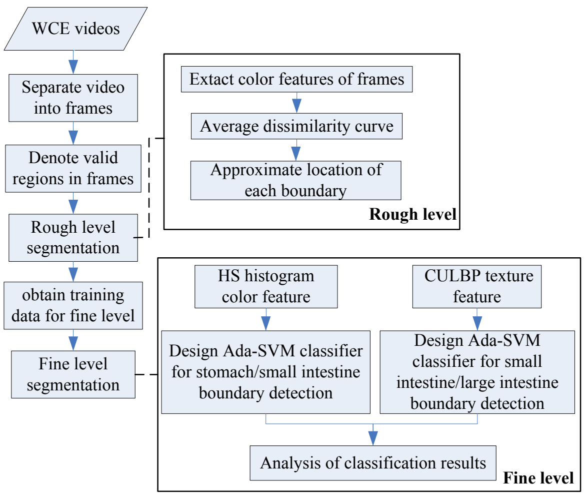


Fig.2 Flowchart of the proposed WCE video segmentation algorithm

texture feature of color uniform local binary pattern (CULBP) and Ada-SVM algorithms are employed. The CULBP feature is robust to the variation of illumination and the Ada-SVM algorithm boosts the performance of general SVM classifier. The experiments demonstrate a promising performance of the proposed approach.

This paper is organized as follows: Section II introduces the method for WCE video segmentation, in which CULBP algorithm and Ada-SVM classifier has been presented in details. Then experiments show the effectiveness and feasibility of the proposed algorithm in Section III. Finally Section IV discusses the conclusion and future work.

## II. METHOD FOR WCE VIDEO AUTOMATIC SEGMENTATION

In our previous work [11], we proposed a two-level approach for WCE video segmentation, and the flowchart of this method is shown in Fig. 1. First, a WCE video is separated into frames, and valid regions are denoted using color feature and wavelet texture feature. Secondly, in the rough level, approximate location of each boundary has been obtained. In this stage, an average dissimilarity curve is presented based on mean and variance color features in Lab color space. Thirdly, frames sampled with an interval of 16 frames between the two approximate boundaries of organs are considered as the training data for the fine level classification. At last, HS histogram feature and ULBP texture are applied to the original SVM classifier to get extract boundaries around the approximate locations.

From the experiments in [11], it shows that the segmentation in small intestine/large intestine performs much worse than that in stomach/small intestine, because it used the same feature for both two boundary detection. However, images in small intestine contain similar color feature with images in large intestine and some images in stomach have few texture information. Thus, in this paper, we design two classifiers using different features for stomach/small intestine boundary detection and small intestine/large intestine boundary detection respectively, shown in Fig. 2. In the stage of stomach/small intestine boundary detection, HS color histogram feature is utilized. And then, a new CULBP texture feature is presented for small intestine/large intestine boundary detection, which contains two different kinds of texture patterns 1) the color norm patterns, and 2) the color angular patterns. The color angular patterns are robust to variation of illumination. Then we proposed a novel Ada-SVM classifier based on the Adaboost algorithm and general SVM classifier. Adaboost algorithm, short for Adaptive Boosting, is a machine learning algorithm which can elevate the performance of a weak classifier by generating a strong classifier out of a set of weak classifiers. However, in this work, we break the general boosting principle, and the experiments show that the Ada-SVM classifier performances even better than the general SVM classifier with RBF kernel.

### *A. HS Color Histogram*

Since some WCE frames in stomach contain few texture features, color distribution is considered as a primary characteristic for stomach/small intestine boundary detection. HSI color space is applied, which decomposes the image into components of chromaticity and luminance. HS color histogram demonstrates robustness to illumination change, and it is reported in [12] that color histogram is invariant to image scale changes, translation and rotation about the viewing axis, and partial occlusion, so the HS color histogram is considered in this paper. First, valid regions in each WCE frame have been denoted by color and wavelet texture features, because invalid regions, which contain useless information, may affect the discrimination of features [11]. However, the valid regions have various shapes, so the pixel number in valid region is diverse in different frames, as shown in Fig. 3. Thus, the normalized HS histogram is carried

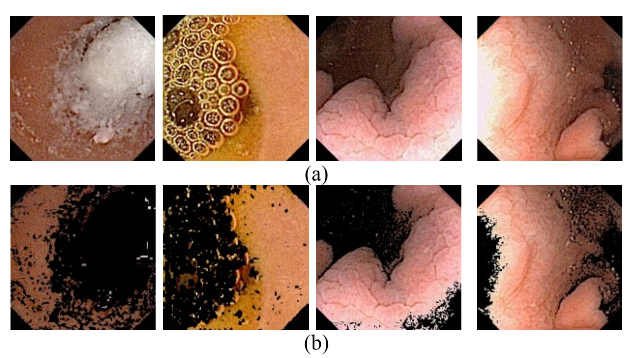


Fig.3. Examples of denoting valid regions in WCE frames. (a) Original WCE frames with gastric juice, bubbles, shadows and excessive bright regions, respectively. (b) Results of denoting invalid regions in images of (a). The black regions are invalid, and the color parts are valid.

out. And we uniformly quantize the component of H and S in each frame into 16 bins respectively.

$$hisH(i) = \frac{n_i^H}{N} \quad (1)$$

$$hisI(i) = \frac{n_i^I}{N} \quad (2)$$

where $N$ is the total number of pixels in valid regions of a frame, $n_i^H$ and $n_i^I$ is the frequency of $i$-th bin in $H$ and $I$ channels of the valid regions, and $i$ = 1, 2... 16.

### B. A Novel Color Uniform Local Binary Pattern

Texture is another important feature in medical image analysis. And the local binary pattern (LBP) operator is one of the most widely used texture features, which is robust to illumination. And color LBP is proposed in order to utilize color information for classification. However, in most previous works, color LBP vector is obtained by concatenating the three texture feature vectors of different color channels together. In this paper, a new method is presented with color norm patterns and color angular patterns to integrate color information in different color channels. In the remainder of this subsection, details about the CULBP method are described.

Supposing $I$ is a RGB WCE frame with size of 288×288 pixels, and each pixel contains a color vector $v = (r, g, b)^T$. Then a color norm value $Vnorm$ is described as follows:

$$Vnorm = \|v\| = \sqrt{r^2 + g^2 + b^2} \quad (3)$$

The color norm value combines 3 different color channels so that the classification performance becomes more effective and reliable than using only one channel information.

Next, the color norm vector is applied to the LBP algorithm to present the color norm patterns. The general LBP is calculated in a circle, and the center pixel is set as the threshold. Then other pixels are compared with the threshold to obtain a binary value. The LBP value of each center pixel is achieved by a weighted summation of other pixels according to its position. $R$ presents the distance of the center pixel and other pixels. $P$ is the number of pixels of the circle except the center one. $LBP_{P,R}$ stands for the LBP process.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (4)$$

where $s(x)$ is a sign function, $g_c$ is the grey value of the center pixel, and $g_p$ is the grey value of other pixels.

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (5)$$

However, it is found that some of the patterns often appear in low frequency and some in high frequency that can be considered as the basic property of texture. The Uniform pattern is proposed to extend LBP in [13], which contains most of the two alterations from 0 to 1 in binary encoding. Here, we use the color norm value to calculate the color uniform LBP.

$$LBP_{P,R}^{u,norm} = \begin{cases} \sum_{p=0}^{P-1} s(Vnorm_p - Vnorm_c) & if\ U(LBP_{P,R}^{norm}) \leq 2 \\ P+1 & otherwise \end{cases} \quad (6)$$

where $U(LBP_{P,R}^{norm})$ is the measurement of color norm uniform pattern, $Vnorm_c$ is the color norm value of the center pixel, and $Vnorm_p$ is the color norm value of other pixels..

$$U(LBP_{P,R}) = \left|s(Vnorm_{p-1} - Vnorm_c) - s(Vnorm_0 - Vnorm)\right| + \sum_{p=1}^{P-1}\left|s(Vnorm_p - Vnorm_c) - s(Vnorm_{p-1} - Vnorm_c)\right| \quad (7)$$

Patterns with $0 < U(LBP_{P,R}^{norm}) \leq 2$ is considered, which are relevant and approximately correspond to edges, line endings, and corners. Then, the histogram of the color norm patterns is calculated, and each color norm pattern vector contains 7 bins.

$$H_{norm} = [h_1, h_2, \dots, h_7]^T \quad (8)$$

In order to extract discriminative color patterns between different color channels, the angle between different color channels at each pixel is calculated. The angle can be described as the ratio of values between different color channels at a pixel, which presents directional information of a color vector and is robust to the variation of illumination.

The angle between different color channels are defined by

$$\tan\alpha_{(r,g)} = \frac{r}{g+\Delta}$$
$$\tan\alpha_{(b,g)} = \frac{b}{g+\Delta} \quad (9)$$
$$\tan\alpha_{(r,b)} = \frac{r}{b+\Delta}$$

where $v = (r, g, b)^T$ is the color vector in a pixel, and $\Delta$ is a very small constant to avoid a zero-valued input in the denominator term. The color angle patterns is described as follows

$$LBP_{P,R}^{u2,color} = \begin{cases} \sum_{p=0}^{P-1} s(\alpha_p - \alpha_c) & if\ \ 0 < U(LBP_{P,R}^{color}) \leq 2 \\ P+1 & otherwise \end{cases} \quad (10)$$

Using (9) and (10), the histograms of the color angle patterns for every pair of color channels are calculated, and each color angle pattern vector contains 3×7 bins [16].

$$H_{(r,g)} = \left[h_{(r,g),1}, h_{(r,g),2}, \dots, h_{(r,g),7}\right]^T$$
$$H_{(b,g)} = \left[h_{(b,g),1}, h_{(b,g),2}, \dots, h_{(b,g),7}\right]^T \quad (11)$$
$$H_{(r,b)} = \left[h_{(r,g),1}, h_{(r,g),2}, \dots, h_{(r,g),7}\right]^T$$

### C. Ada-SVM Algorithm

SVM is primarily a classifier method that performs classification tasks by constructing hyper-planes in a multidimensional space that separate cases of different class labels [14]. The RBF kernel is the most popular choice of kernel types used in SVM by far because of its localized and finite response across the entire range of the real x-axis. Thus, RBFSVM is selected in our process.

In order to increase the performance of SVM classifier, the Ada-SVM classifier is proposed in this paper, which uses SVM as component of the Adaboost classifier [15]. The AdaBoost (Adaptive Boosting) algorithm is proposed in 1995 by Yoav Freund and Robert Shapire as a general method for generating a strong classifier out of a set of weak classifiers, which boosts the performance of the classifier. It adjusts the weight of each weak classifier and the distribution of training samples at each loop according to its accuracy in previous loop. In this paper, we break the general boosting principle, since adjusting the parameter gamma can obtain a set of moderately accurate RBFSVMs for AdaBoost. Moreover, the adaboost algorithm forces some SVM components focus on the misclassification samples, which boosts the performance of SVM classifier.

**1. Input**: Training Data with labels

$$A = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$$

**2. Initialize:**

a. Calculate the weights of training Data:

$$w_i^1 = \frac{1}{N},\ i = 1, 2, \ldots, N$$

b. Obtain a training set of M samples from the training data in A, randomly.

**3. Do while** (t < TMAX)

a. If $t > 1$, re-sample the training data to obtain a new training set according to the weights.

b. Train a RBF-SVM classifier $h_t$ on the sampled training set, where crossing validation is applied to obtain the gamma and C parameters.

c. Calculate the training error of $h_t$ on the whole training data $\varepsilon_t = \sum_{i=1}^{N} w_i^t,\ y_i \neq h_i(x_i)$

d. If $\varepsilon_t > 0.5$, break;

e. Set the weight of classifier $h_t$: $\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$

f. Update the weights of the whole training data, $w_i^{t+1} = \frac{w_i^t \exp(-\alpha_t y_i h_t(x_i))}{C_t}$, where $C_t$ is a normalization constant.

g. t++;

**4. Output:** $f(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x))$

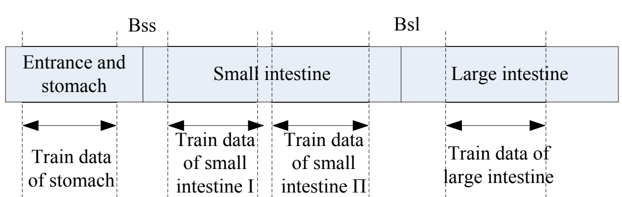


Fig.4 Train data collection in fine level. Bss is the approximate boundary of stomach/small intestine and Bsl is that of small intestine/large intestine. Train data of stomach and train data of small intestine I is collected for stomach/small intestine classifier. Train data of small intestine Π and train data of Train data of large intestine is collected for small intestine/large intestine classifier.

*D. Analysis of the Classification Results*

In order to obtain the extract boundaries between two adjacent organs from the classification results, we define the boundary as the point which makes the classification error of both organ regions get the minimum. Suppose that *Predict* is the result sequence, $Error_1$ is the classification error of the former organ region and $Error_2$ is error of the later organ region.

$$\begin{aligned} Error_1(i) &= \begin{cases} 0 & if\ Predict(i) = 1 \\ 1 & if\ Predict(i) = -1 \end{cases} \\ Error_2(i) &= \begin{cases} 0 & if\ Predict(i) = -1 \\ 1 & if\ Predict(i) = 1 \end{cases} \end{aligned} \quad (12)$$

$$Predict = \{Predict(i) \mid Predict(i) \in [1, -1], i = 1, 2, \ldots, Len\}$$

where *Len* is the length of the result sequence.

Then, the whole error of the result sequence *E* can be calculated

$$E(k) = \sum_{i=1}^{k} Error_1(i) + \sum_{i=k+1}^{Len} Error_2(i) \quad (13)$$

The minimum of the error sequence *E* is considered as the separation point of the two regions of organ.

## III. EXPERIMENTS

*A. Dataset and Experiment Design*

The Prince of Wales Hospital in Hong Kong provides the WCE video data used in these experiments. There are three videos lasting on an average of 8 hours, and contain 61427, 57694 and 60170 with 288 × 288 pixels of frames respectively. The camera in WCE capsule captures 2 frames per second. And the exact boundaries of each video are obtained by experts.

In our experiments, approximate boundaries are obtained from the average dissimilarity curve with simple color features in the rough level segmentation. And the boundaries are located with the maximum error of 2 minutes. Define $B_{ss}$ as the approximate boundary of stomach/small intestine and $B_{sl}$ as that of small intestine/large intestine. Training data of stomach and small intestine I are collected for stomach/small intestine classifier, with label 0 and 1 respectively. As shown in Fig. 4, training data of small intestine II and large intestine are collected for small intestine/large intestine classifier, with label 1 and 0 respectively. And the training data of stomach is sampled from $B_{ss}$-40 minutes to $B_{ss}$-10 minutes in the video, the training data of small intestine I is from $B_{ss}$+10 minutes to $B_{ss}$+40 minutes, the training data of small intestine II is from $B_{sl}$-40 minutes to $B_{sl}$ -10 minutes and training data of large intestine is from $B_{sl}$+10 minutes to $B_{sl}$+40 minutes. All training frames are sampled with an interval of 8 frames in the video, and there are 450 positive training samples and 450 negative training samples for each classifier. Then, 960 frames ( minutes in the video) around the each approximate boundary are selected for classification to locate the exact boundary of each organ.

Three experiments are designed to demonstrate the reliability and efficiency of our proposed algorithm. The first experiment shows the effectiveness of the proposed feature, which compares the classification performance of using CULBP feature with that of original LBP feature in locating the boundary between small intestine and large intestine. In the second experiment, an analysis of the proposed Ada-SVM algorithm is presented and we compare the performance of Ada-SVM algorithm with SVM classifier. The Ada-SVM classifier appears to be more robust to parameter change. And in the last experiment, we test the performance of segmentation algorithm in [6] and [11]. In [6], MPEG-7 color feature is used and a global model fitting approach is applied to the classification results of SVM classifier. But it's a time consuming process of estimating and judging all frames in a WCE video.

The accuracy of the video segmentation algorithms is assessed as the error frames between the boundary obtained from the experiments and the one manually labeled by an expert. The mean and the median errors are considered in experiment results. The mean error is the average error of all test videos and the median error is the middle error value. In our experiment, precision and recall are also utilized to analyze the classification performance of classifiers. Precision indicates the fraction of the positives detected that are actually correct. Recall indicates the probability of correctly detecting a positive test sample and is independent of class priors. In our program, frames in small intestine are considered as positive samples which will use for further process (such as lesion detection). And mistaking classifications of true positive frames may cause fatal result. Thus, recall is a more important indicator for the WCE video segmentation. And both recall and precision should be as high as possible for video segmentation.

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

where $TP$ is the number of true positive frames, $FP$ is the number of false positive frames, $TN$ is the number of true negative frames, and $FN$ is the number of false negative frames.

### B. Experiment Results

The result of the first experiment is shown in Table I, which compares the performance of the proposed CULBP feature and general ULBP feature. SVM classifier with RBF kernel is used in this experiment. From the result, it can be seen that CULBP feature reaches an average of 88.19% accuracy and 96.91% recall, which performs better than general ULBP with an average of 87.64% accuracy and 80.67% recall. The reason is that the proposed CULBP feature combines texture information of three color channels and it contains not only color value but also the angle information between different color channels, which is more robust to the variation of illumination and more discriminative for classification.

In the second experiment, we test the performance of the proposed Ada-SVM classifier, and a comparison between Ada-SVM classifier and original classifier with RBF kernel has been shown in Table II and III. In the stomach and small intestine classification, the precision of Ada-SVM classifier has a slight decrease (91.37%), but still satisfactory, than that of original SVM classifier (93.76%). However, the recall of Ada-SVM classifier, which is a more important indicator, has a great improvement of 9.51%. In the small intestine and large intestine classification, both precision and recall of the proposed Ada-SVM classifier are higher than that of original SVM classifier, reaching 90.35% and 97.28% respectively, because the Ada-SVM classifier integrates Adaboost algorithm with SVM classifier. The Adaboost algorithm can "boost" the performance of the general SVM classifier, which adjusts the weights of training data for each weak classifier in the loops and it just requires the weak classifier to achieve accuracy higher than 50%. At the same time, it solves the problem that the SVM classifier is sensitive to parameter change.

Tables IV and V show the performance of different WCE video segmentation methods. We divide this experiment into two stages. First, we compare the precision and recall of different methods (see Table IV). In order to compare the classifier results, same testing data are applied to the three methods. 960 frames around each approximate boundary are collected for testing. There are totally 5760 testing frames. The average recall of method in [6] just receives 81.92% and 75.17% in stomach/small intestine classification and small intestine /large intestine classification respectively, which appears lower than that of the other two methods. This can be explained by that it only utilizes the color features, in which the texture feature isn't considered. However, texture features in small intestine and large intestine present much more discriminative than color features. Second, the segmentation performance of WCE video has been tested (see Table V). From the table, it can be seen that method in [6] has the worst performance with average errors of 2563 and 3136 frames in stomach/small intestine boundary detection and large/small intestine boundary detection respectively, which even cannot be accepted. Because, in [6], all frames in a test video need to be classified and it requires the diversity of training data, but in our experiment there are only three entire WCE video. Then, the training data for method in [6] is selected from the rest two WCE videos, except the test one. Thus, it cannot achieve a reasonable performance due to the lack of diverse training data. However, both training data of method in [11] and the proposed method are obtained from the test video itself around the approximate boundary gotten from the rough level, so they can get a much better performance. It also can be seen form Table V that the proposed method has a great improvement in small intestine and large intestine segmentation with average errors of 7 and 23 frames, which can be explained by that we design two classifiers for two different organ boundaries detection and the proposed CULBP feature contains both texture and color information. Moreover, the Ada-SVM classifier is more robust than the general SVM classifier used in [11].

TABLE I. COMPARISON OF CULBP AND ORIGINAL ULBP

| | CULBP | | Original ULBP | |
|---|---|---|---|---|
| | Precision | Recall | Precision | Recall |
| Video 1 | 95.60% | 99.67% | 95.33% | 89.29% |
| video 2 | 70.18% | 94.62% | 70.74% | 92.04% |
| Video 3 | 98.79% | 96.44% | 96.85% | 60.67% |
| Average | **88.19%** | **96.91%** | **87.64%** | **80.67%** |

TABLE II. PERFORMANCE OF ADA-SVM CLASSIFIER

| | Stomach/Small intestine | | Small intestine/Large intestine | |
|---|---|---|---|---|
| | Precision | Recall | Precision | Recall |
| Video 1 | 93.51% | 87.63% | 95.79% | 99.34% |
| ideo 2 | 89.61% | 88.68% | 75.86% | 95.27% |
| Video 3 | 90.99% | 89.18% | 99.39% | 97.23% |
| Average | **91.37%** | **88.50%** | **90.35%** | **97.28%** |

TABLE III. PERFORMANCE OF ORIGINAL SVM CLASSIFIER

| | Stomach/Small intestine | | Small intestine/Large intestine | |
|---|---|---|---|---|
| | Precision | Recall | Precision | Recall |
| Video 1 | 90.28% | 77.28% | 95.60% | 99.67% |
| video 2 | 98.39% | 88.07% | 70.18% | 94.62% |
| Video 3 | 92.60% | 74.61% | 98.79% | 96.44% |
| Average | **93.76%** | **79.99%** | **88.19%** | **96.91%** |

TABLE IV. CLASSIFICATION PERFORMANCE OF DIFFERENT APPROACHES

| | Stomach/Small intestine | | Small intestine/Large intestine | |
|---|---|---|---|---|
| | Average precision | Average recall | Average precision | Average recall |
| Method in [6] | 79.34% | 81.92% | 93.98% | 75.17% |
| Method in [11] | 90.26% | 87.99% | 86.23% | 92.77% |
| Proposed method | **91.37%** | **88.50%** | **90.35%** | **97.28%** |

TABLE V. SEGMENTATION ERRORS OF DIFFERENT APPROACHES

| | ErrorSS | | ErrorSL | |
|---|---|---|---|---|
| | Mean | Median | Mean | Median |
| Method in [6] | 2563 | 112 | 3136 | 2563 |
| Method in [11] | 10 | 19 | 60 | 87 |
| Proposed method | **7** | **7** | **23** | **5** |

ErrorSS: errors of the boundary between stomach and small intestine.
ErrorSL: errors of the boundary between small intestine and large intestine.

## IV. CONCLUSION

In this paper, a novel effective approach has been proposed for WCE video segmentation based on color uniform LBP and boosting SVM algorithm. There are three contributions in this paper. First, a two-level segmentation method is utilized for segmentation which reduces the computation time and overcome the difficulty of training data acquisition. Second, we propose the CULBP feature to advance the discrimination of general uniform LBP feature, which includes two kinds of color patterns, color norm patterns and color angle patterns. It is more robust to variation of illumination and more discriminative for classification. The last one is that an Ada-SVM classifier is presented in this paper, which is non-sensitive to changes of SVM classifier parameters. The experiments indicate that the average precision and recall achieve as high as 90.86% and 92.89% respectively, and the mean segmentation error is less than 23 frames. In the future, we will have more WCE videos in our experiments and investigate new methods for abnormality detection in small intestine.

## REFERENCES


[1] A. Karargyris and N. Bourbakis, "A methodology for detecting bloodbased abnormalities in wireless capsule endoscopy videos," *in Proc. 8th IEEE Int. Conf. BioInf. BioEng.*, Oct. 8–10, 2008, pp. 1–6.

[2] D. G. Adeler, C. J. Gostout, "Wireless capsule endoscopy," *Hospital Physician*, May (2003) 14-22.

[3] S. Iogagiu, L. Ciobanu, O. Pascu, "Colon Capsule endoscopy: a new method of investigating the large bowel," *Journal of Gastrointestinal and Liver Diseases*, vol. 17, pp. 347-352, 2008.

[4] J. Berens, M. Mackiewicz, D. Bell, "Stomach, intestine and colon tissuediscriminators for wireless capsule endoscopy images," *Proc. of SPIE Conference on Medical Imaging*, vol.5747, Bellingham, WA, 2005, pp. 283-290.

[5] J. Lee, J. Oh, S.K. Shah, X. Yuan and S.J. Tang, "Automatic classification of digestive organs in Wireless Capsule Endoscopy videos," *SAC '07: Proc. of the 2007 ACM symposium on Applied computing*, pp. 1041-1045, 2007

[6] J.P. Silva Cunha, M. Coimbra, P. Campos, J.M. Soares, "Automated topographic segmentation and transit time estimation in endoscopic capsule exams," *IEEE Trans. Medical Imaging*, vol. 4, pp. 570–578, July 1993.

[7] M. Mackiewicz, J. Berens and M. Fisher, "Wireless capsule endoscopy color video segmentation," *IEEE Trans. Medical Imaging*, vol. 27, pp. 1769- 1781, December 2008.

[8] B. Li, Max Q.-H. Meng, "Capsule endoscopy video boundary detection," *Proc. of the IEEE, Int. Conf. on Information and Automation*. June 2011.

[9] B. Li, Max Q.-H. Meng, C. Hu, "Motion analysis for capsule endoscopy video segmentation," *Proc. of the IEEE, Int. Conf. on Automation and Logistics*, August 2011.

[10] Y. Shen, P. Guturu, and Bill P. Buckles, "Wireless capsule endoscopy video segmentation using an unsupervised learning approach based on probabilistic latent semantic analysis with scale invariant features", *IEEE Trans. Technology in Biomedicine*, vol. 16, no. 1, pp. 98-105, January 2012.

[11] R. Zhou, B. Li, Z. Sun, C. Hu, Max Q.-H Meng, "Wireless capsule endoscopy video automatic segmentation," *Proc. of the IEEE, Int. Conf. on Robotics and Biomimetics*, December 2012.

[12] M. Swain and D. Ballard, "Color indexing," *Int. J. Comput.* Vis., vol. 7, pp. 11–32, 1991.

[13] G. Sharma, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, July 2002.

[14] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, 1995.

[15] R. E. Schapire and Y. Freund, "Boosting the margin: a new explanation for the effectiveness of voting methods," *The Annals of Statistics*, vol. 26(5), pp. 1651–1686.

[16] S.H. Lee, J.Y. Choi, Y. M. Ro, K, N. Plataniotis, "Local color vector binary patterns from multichannel face image for face recognition," *IEEE Transactions on Image Processing*, pp. 2347-2353, vol.21, no.4, 2012.