33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Bleeding Detection in Wireless Capsule Endoscopy Images Based on Color Invariants and Spatial Pyramids Using Support Vector Machines

Guolan Lv, Guozheng Yan, and Zhiwu Wang

Abstract— **Wireless capsule endoscopy (WCE) is a revolutionary imaging technique that enables detailed inspection of the interior of the whole gastrointestinal tract in a non-invasive way. However, viewing WCE videos is a very time-consuming, and labor intensive task for physicians. In this paper, we propose an automatic method for bleeding detection in WCE images. A novel series of descriptors which combine color and spatial information is designed in a way that local and global features are also incorporated together. And a kernel based classification method using histogram intersection or chi-square is deployed to verify the performance of the proposed descriptors. Experiments demonstrate that the proposed kernel based scheme is very effective in detecting bleeding patterns of WCE images.**

I. INTRODUCTION

Gastrointestinal (GI) diseases have been afflicting large proportions of people all around the world for years. However, the small intestine, as the longest portion of the digestive tract, was technically difficult to be examined by conventional diagnostic procedures, such as push enteroscope, intraoperative enteroscopy and radiographic techniques [1]. Most recently, wireless capsule endoscopy (WCE), which enabled the non-invasive inspection of the whole small bowel [2], showed superior performance in the diagnosis of small bowel pathology, especially in the area of obscure GI bleeding (OGIB) [1].

WCE is firstly swallowed by a patient, and then moves by peristalsis. It captures 2 images per second and generates about 55,000 images during the 8 hour's examination. Finally, the recorded images are downloaded to a work station, and reviewed by a physician. However, the reading and interpretation time is 40-60 min on average [3]. The time-consuming task has become one of the major disadvantages of WCE. Therefore, automatic bleeding detection of WCE images is an urgent need.

Recently, many efforts have been dedicated toward the automatic detection of bleeding patterns in WCE images, since bleeding is a very significant symptom of GI diseases and most of the causes of bleeding can be cured or controlled. The Suspected Blood Indicator (SBI), designed for the detection of OGIB, was provided by Given Imaging. However, the sensitivity and specificity of this tool were reported to be unsatisfactory [4]. In [5], the author proposed an unsupervised method using Expectation Maximization (EM) clustering algorithm to automatically detect the bleeding regions. Unfortunately, color ranges of the blood pixels and the non-blood pixels overlapped and varied due to the changing imaging environment in GI tract. The author in [6] exploited a supervised method utilizing multilayer perceptron neural network (MLP). Considering the illumination variation, chrominance moments, an invariant local color feature in HSI color space, was extracted. However, the method only showed the classification sensitivity and specificity of image patches based on local color texture features, and the overall classification performance based on image levels was not clear.

In our study, a series of histogram representations of WCE images are proposed. The novel descriptors incorporate color and spatial information by combining illumination invariant color histograms and spatial pyramids. Therefore, they are robust to light variation within GI tract and invariant to translation and rotation. Two kinds of photometric invariant color histograms (hue histogram and transformed color histogram) are explored. A kernel based classification method using histogram intersection kernel or chi-square kernel which is especially suitable for histogram features is proposed. Experimental results show that our method is a promising diagnostic proposal and hence lead to wider acceptance of WCE.

Our method is mainly inspired by the following research: (1) the spatial pyramid representation of images proposed in [7], and (2) the color invariant descriptors discussed in [8].

The rest of the paper is organized as follows. Section II describes the details of the descriptors, termed PCIH (Pyramid of Color Invariant Histograms), and the idea is illustrated in Fig. 1. Section III explains the kernel based SVM classifier. Experiments and results are then discussed in Section IV. Finally, in Section V, conclusions are drawn.

II. FEATURE EXTRACTION

Bleeding detection in WCE videos must take into consideration the following issues:

1) Bleeding patterns manifest themselves by red colors different from colors of normal region. Intuitively, color appearance is also the first and most important basis for physicians to make diagnostic decisions.

Manuscript received April 15, 2011. This work was supported by the National Hi-Tech Research and Development Program (863) of China (No. 2006AA04Z368) and the National Natural Science Foundation of China (No. 30570485).

Guolan Lv is with the School of Electronics, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. (corresponding author to provide phone: 86-021-34204435; e-mail: 1090359015@sjtu.edu.cn).

Guozheng Yan and Zhiwu Wang are with the School of Electronics, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. (e-mail: gzhyan@sjtu.edu.cn, zwwang@sjtu.edu.cn).

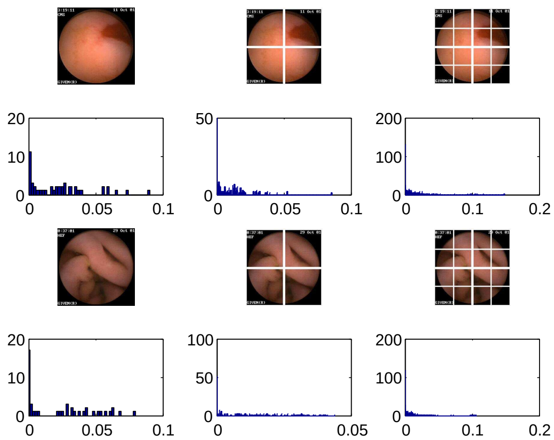

Fig. 1. Pyramid of transformed color histogram representation. The first row: a bleeding image and grids for levels $l = 0$ to $l = 2$; below: histogram representations for each level. The next two rows are for a non-bleeding image.

2) Light intensity in digestive tract is time-varying, and changes in the illumination can greatly affect the performance of bleeding detection if the descriptors used are not robust to these variations.

3) The shape and size of bleeding regions are uncertain. Some are large enough to cover the whole image, while others are only spots. So a combination of both global and local feature is better than global or local feature alone. And spatial information should also be considered.

A. Color Invariant Descriptor

Color histogram is an extremely simple way of representing the color distribution in an image which is suitable for real-time computer aided systems. For digital images, histogram is obtained by discretization of the image colors into a number of bins, and counting the number of times each color occurs in each bin. Histograms are invariant to translation and rotation, and vary slowly with view angle [9].

However, color histogram is highly sensitive to lighting intensity changes. Therefore, to increase the photometric invariance and the discriminative power, several color histograms were studied [10]. We choose Hue Histogram and Transformed Color Histogram which are both invariant to light intensity change and shift. They are described as follows:

1) Hue Histogram: In HSI color space, hue is known to be invariant with respect to lighting geometry and specularities [10]. However, hue is unstable at the achromatic axis ($R = G = B$). As a result, a small perturbation of RGB values might lead to a large jump in the hue values [11]. It was found that the certainty of hue is inversely proportional to the saturation. Therefore hue histogram can be more robust by weighting each sample by its saturation [10]. The robust hue histogram is invariant with respect to light intensity change and shift [8].

2) Transformed Color Histogram: RGB histogram itself is sensitive to photometric variations. However, scale-invariance and shift-invariance with respect to light intensity can be achieved by normalizing the pixel value distribution as follows:

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} \frac{R-\mu_R}{\sigma_R} \\ \frac{G-\mu_G}{\sigma_G} \\ \frac{B-\mu_B}{\sigma_B} \end{pmatrix} \tag{1}$$

where μ_X is the mean and σ_X is the standard deviation of the distribution in color channel $X(X = R, G, B)$ over the area under consideration [10].

B. Pyramid of Color Invariant Histogram - PCIH

The major drawback of color histogram representation is that it discards all the spatial information of images, which is a key attribute for image classification tasks. To overcome this limitation, we introduce spatial information and create the multi-resolution color histogram using the spatial pyramid histogram approach proposed by Lazebnik in [7]. In this paper, a similar scheme is designed, consisting of color invariant histograms over each image subregions at each resolution level - Pyramid of Color Invariant Histograms (PCIH), including Pyramid of Hue Histograms (PHH), and Pyramid of Transformed Color Histograms (PTCH). The new descriptors have the advantage of combing global and local features so that small bleeding regions will not be misclassified.

The process of building PCIH is as follows: for a spatial pyramid histogram with L levels, the grid at level l has 2^l cells along each dimension of images. The level 0 histogram is first generated over the entire image and represented by a vector of length M (M is the histogram bin size). For level l, the image is divided into 4^l equal sized sub-regions and a level l histogram is computed over each region and represented by a $4^l \times M$ dimensional vector. The process is repeated by subdividing each image into increasingly finer spatial grid and computing histograms in each region until level L is reached. Finally, the appropriately weighted histograms at all resolutions are concatenated, and the PCIH descriptor of the entire image is formed to be a vector with dimensionality $M \sum_{l=0}^{L} 4^l$. For example, if we describe a WCE image with $L = 3$ level and $M = 36$ bins, the resulting vector has length 756.

$$\begin{aligned} H'_0 &= \frac{1}{2^L} H_0, l = 0 \\ H'_l &= \frac{1}{2^{L-l+1}} H_l, l > 0 \\ H &= (H'_0, H'_1, \dots, H'_L) \end{aligned} \tag{2}$$

where $H'_l, l = 0, 1, \dots, L$ represents the illumination invariant color histogram of level l, H'_l represents the weighted histogram of level l, and H is the vector form of PCIH.

Note that histograms at finer resolutions are given more weight, because small bleeding regions are more likely to be misclassified. Fig.1. shows that bleeding images have a similar PCIH representation and that this representation differs from non-bleeding images.

III. SUPPORT VECTOR MACHINES

Support Vector machines (SVMs) is a kernel-based machine learning technique which has been widely used in real-world classification problems in various domains. Due to its strong theoretical foundation, good generalization capability, and ability to find global classification solutions, SVMs is usually preferred by many researchers over other classification paradigms.

Given a binary classification problem: $\{(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)\}$, where $x_i \in \Re_n$ represents an n-dimensional data point, and $y_i \in \{-1, 1\}$ represents the corresponding class. The support vector machines require the solution of the following optimization problem:

$$\begin{aligned} &\min_{\omega,b,\varepsilon} \left(\frac{1}{2}\omega \cdot \omega + C \sum_{i=1}^{k} \varepsilon_i \right) \\ \text{subject to } & y_i(\omega \cdot \phi(x_i) + b) \geq 1 - \varepsilon_i \\ & \varepsilon_i \geq 0, i = 1, \ldots, k \end{aligned} \tag{3}$$

where ε_i is the slack variable which holds for misclassified example, and C is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is called the kernel function.

In this section, we will introduce two nonlinear Mercer kernels for SVM that are found to be most suitable on histogram representations.

A. Histogram Intersection Kernel

Histogram intersection proposed in [9] is an efficient way of matching histograms, and has been proven to be a useful kernel function in histogram based image classification problems. It counts the number of pixels that fall in the same bins in the two histograms h^1 and h^2, and is defined in the following equation:

$$K(h^1, h^2) = \sum_{i=1}^{n} \min(h_i^1, h_i^2) \tag{4}$$

Histogram intersection is extremely simple and easy to implement.

B. Chi-Square Kernel

The chi-square kernel is derived from the chi-square distribution. It is a simple and effective similarity measure for histogram comparison. The chi-square kernels are defined as follows:

$$K^2(h^1, h^2) = 1 - \sum_{i=1}^{n} \frac{(h_i^1 - h_i^2)^2}{\frac{1}{2}(h_i^1 + h_i^2)} \tag{5}$$

IV. EXPERIMENTS

Nine videos of different bleeding types were obtained from Given Imaging (http://www.capsuleendoscopy.org/). Each video is 20 seconds long. Since WCE images were collected at 20 frames/s and neighboring frames varied slightly, images 256×256 were selected at an interval of greater than or equal to three images depending on their variations. Also images that are too dark or contain many visual contaminations were removed. Finally, a representative subset of 560 examples composing of 280 bleeding images and 280 non-bleeding images was obtained from all the videos. The dataset was split randomly into 1:1 proportion for training and testing, respectively.

To evaluate the performance of the proposed algorithms, the widely used libsvm software package [12] was selected. We considered the histogram intersection kernel and chi-square kernel as precomputed kernels for SVM training and selected the optimal value for parameter C over the range: $log_2 C = \{1, 2, \ldots, 15\}$. We evaluate the performance of a model trained on each C by using five-fold cross-validation result on the training dataset. After finding the optimal parameter, a new SVM model was trained by using the complete training dataset on the parameter and the performance of that model was tested on the remaining testing partition.

For performance evaluation, sensitivity, specificity, accuracy and error rate, widely used in medical diagnostics, were investigated in the experiment.

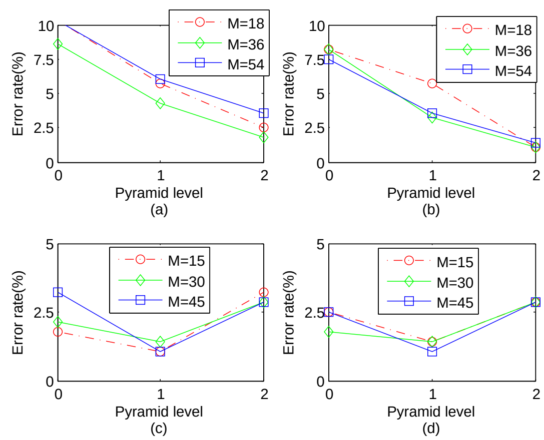

Fig. 2. (a), (b) shows the performance for PHH using chi-square kernel and histogram intersection kernel respectively; (c), (d) shows the performance for PTCH using chi-square kernel and histogram intersection kernel respectively.

Fig. 2 shows that if we change the number of histogram bins M in a range $[18\ 36\ 54]$ for hue histogram and $[15\ 30\ 45]$ for transformed color histogram, the optimal results are obtained at $M = 36$ for PHH, and $M = 15$ for PTCH. Note that the classification performance is not very sensitive to the number of histogram bins.

For PHH, results improve dramatically as pyramid level increase from $l = 0$ to $l = 2$. For PTCH, the performance is optimal at $l = 1$. This means that color and spatial

Fig. 3. A comparison of performance for PHH descriptors using chi-square kernel and histogram intersection kernel.

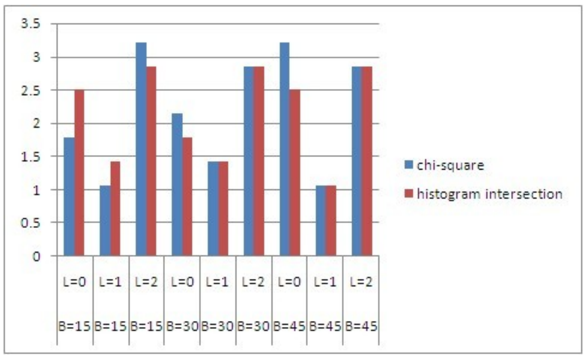

Fig. 4. A comparison of performance for PTCH descriptors using chi-square kernel and histogram intersection kernel.

information fusion exhibits a statistically significant benefit. However, the performance for PTCH drops at the highest level $l = 2$. This indicates that at the higher levels regions are smaller and feature lengths are much longer, and the global information contributes less to the overall classification performance.

Fig. 3 and Fig. 4 show that the most suitable kernel for PHH is histogram intersection, while the choice of the optimal kernel for PTCH depends on specific conditions. Both kernels can achieve satisfactory results.

Finally, TABLE I compares the classification performances of PHH and PTCH using three different kernels. Obviously, chi-square kernel and histogram intersection kernel perform better than Radial Basis Function (RBF) for the proposed features, and PTCH outperforms PHH.

TABLE I
CLASSIFICATION PERFORMANCES OF PHH AND PTCH USING THREE DIFFERENT KERNELS

	Kernel	Average Accuracy	Average Sensitivity	Average Specificity
PHH	Chi-square	94.1%	94.5%	93.7%
	Histogram intersection	95.6%	96.0%	94.6%
	RBF	89.6%	96.3%	82.9%
PTCH	Chi-square	97.8%	97.7%	97.9%
	Histogram intersection	97.9%	97.8%	98.0%
	RBF	96.6%	97.9%	95.4%

V. CONCLUSIONS

We have introduced a new series of descriptors PCIH (including PHH and PTCH), which incorporate color and spatial information of WCE images, and are robust to illumination variation. The combination of spatial pyramids and robust hue histogram increases accuracy by about 8%, and the combination of spatial pyramids and transformed color histograms achieves about 1% improvement (compared to single level histograms). This demonstrates that color and spatial information fusion is significant for bleeding detection in WCE images. Local and Global feature combined together also outperforms either one. The performance of SVM classifier using histogram intersection and chi-square kernel are both effective to detect bleeding patterns. Therefore, proper selection of numbers of histogram bins, pyramid levels and SVM kernels are all important for good performances. To sum up, the proposed scheme is very effective for bleeding detection. In the future, our work will extend to detecting other diseases in GI tract such as tumor and ulcer.

VI. ACKNOWLEDGMENTS

This work was funded by the National Hi-Tech Research and Development Program (863) of China (No. 2006AA04Z368). Funding Support was also received from the National Natural Science Foundation of China (No. 30570485).

REFERENCES

[1] Sidhu, R., Sanders, D.S., McAlidon, M.E., “Gastrointestinal capsule endoscopy: from tertiary entres to primary care”, *BMJ*, vol. 332, 2006, pp. 528-531.

[2] Iddan G, Meron G, Glukhovsky A, Swain P., “Wireless capsule endoscopy”, *Nature*, vol. 405, 2000, pp. 417.

[3] Pennazio M., “Capsule endoscopy: where are we after 6 years of clinical use?” *Digestive and Liver Disease*, vol. 38, 2006, pp. 867-878.

[4] C. Signorelli, F. Villa, E. Rondonotti, C. Abbiati, G. Beccari, and R. de Franchis, “Sensitivity and specificity of the suspected blood identification system in video capsule enteroscopy”, *Endoscopy*, vol. 37, 2005, pp. 1170-1173.

[5] S. Hwang, J.H. Oh, J. Cox, S.J. Tang, H.F. Tibbals, “Blood detection in wireless capsule endoscopy using expectation maximization clustering”, *Proceedings of SPIE on Medical Imaging*, vol. 6144, 2006, pp. 577-587.

[6] B. Li, Max Q.-H. Meng, “Computer aided detection of bleeding regions in capsule endoscopy images”, *IEEE Transactions on Biomedical Engineering*, vol. 56, 2009, pp. 1032-1039.

[7] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories”, *In Proc. CVPR*, 2006, pp. 2169-2178.

[8] Koen E.A. van de Sande, Theo Gevers, Cees G.M. Snoek, “Evaluating color descriptors for object and scene recognition”, *PAMI*, vol. 32, 2010, pp. 1582-1596.

[9] Michael J. Swain and Dana H. Ballard, “Indexing via color histograms”, *In Proc. ICCV*, 1990, pp. 390-393.

[10] Joost van de Weijer, Cordelia Schmid, “Coloring local feature extraction”, *In Proc. ECCV*, 2006, pp. 334-348.

[11] Theo Gevers, Arnold W.M. Smeulder, “Color based object recognition”, *Pattern Recognition*, vol. 32, 1999, pp. 453-464.

[12] Chih-Chung Chang and Chih-Jen Lin, “LIBSVM: a library for support vector machines”, 2001, Software available available at: http://www.csie.ntu.edu.tw/ cjlin/libsvm