31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

A Comparative Study of Shape Features for Polyp Detection in Wireless Capsule Endoscopy Images

Baopu Li, Max Q.-H. Meng and Lisheng Xu

Abstract—**Wireless capsule endoscopy (WCE) has been gradually employed in hospitals because it can directly view the entire small bowel of a human body for the first time. However, a troublesome problem related to this new technology is that too many images produced by WCE will take a lot of efforts for doctors to inspect. In this paper, we propose a comparative study of shape features aiming for intestinal polyp detection for WCE images. As polyps exhibit strong shape characteristics, also a powerful clue used by physicians, we investigate two kinds of shape features, MEPG-7 region-based shape descriptor and Zernike moments, in our study. With multi-layer perceptron neural network as the classifier, experiments on our present image data show that it is promising to employ both Zernike moments and MEPG-7 region-based shape descriptor as the shape features to recognize the intestinal polyp regions, and a better performance is obtained by the Zernike moments based shape features.**

I. Introduction

DISEASES related to gastrointestinal (GI) tract greatly threaten human's health now. It has been reported that colorectal cancer has been the second leading cause of cancer-related deaths in U.S. [1]. Many diseases in the GI tract can be prevented and cured if early detection is possible. The traditional detection methods such as endoscopy, ultrasound, and computed tomography (CT) scan have demonstrated great values in diagnosing diseases of the digest tract. However, the main body of the GI tract, small intestine, cannot be reached by the traditional endoscopies due to their limitations. In 2000, a new kind of GI endoscopy, i.e., wireless capsule endoscopy (WCE), was created. This new technology of endoscopy, developed by Given Imaging corporation in Israel, almost revolutionize the diagnosis methodology for the digestive tract since this small device can directly view the entire small intestine without pain, sedation, or air insufflation for the first time, and these advantages make it rapidly used in most hospitals to detect the status of the GI tract.

B.P. Li and Max Meng are with the Department of Electronic Engineering at The Chinese University of Hong Kong, Hong Kong SAR, China, email: {bpli&max}@ee.cuhk.edu.hk; L.S.Xu is now is with Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang, 110004 China. (xuls@bmie.neu.edu.cn)

This research is supported by SHIAE project #8115021 of the Shun Hing Institute of Advanced Engineering of The Chinese University of Hong Kong, awarded to Max Meng.

As shown in Fig.1 and Fig.2, wireless capsule endoscopy, measuring 11mm $\times$ 26mm, is a pill-shaped device which consists of some miniature components. After a WCE is ingested by a patient who has a diet for about 12 hours, this small device starts to work and record images while moving forward along the digestive tract. Meanwhile, the images recorded by the camera are sent out wirelessly to a special recorder attached to the waist. This process continues for about eight hours until the WCE battery ends. Finally, all the image data stored in the recorder are downloaded into a computer, and physicians can view the images and analyze potential sources of different diseases in the GI tract. The WCE was approved by U.S. Food and Drug Administration (FDA) in 2001, and it has been reported that this new technology shows great value in evaluating GI bleeding, Crohn's disease, ulcer and other diseases located in the digestive tract [2].

Fig.1 Wireless capsule endoscopy

The understanding of WCE's video data is undertaken by a trained physician. However, it takes a long period of time for physicians to inspect the large number of images it produced. In fact, there are about 50,000 images in total per examination for one patient, and it costs an experienced clinician about two hours on average to review and analyze all the video data [2]. Moreover, abnormalities in the GI tract may be present in only one or two frames of the video, so they might be missed by physicians due to oversight sometimes. Besides, there may be some abnormalities that cannot be detected by the naked eyes due to the great variation in their size, color, texture and distribution. In addition, different physicians may have different findings based on the same image data. All these problems motivate the researchers to develop reliable and uniform assisting approaches to reduce the great burden of the physicians. However, it should be admitted that this goal is very challenging because the true features associated with the diseases are not exactly known. Moreover, different diseases have totally different symptoms in the digestive tract. Even the same disease shows much difference in color and shape.

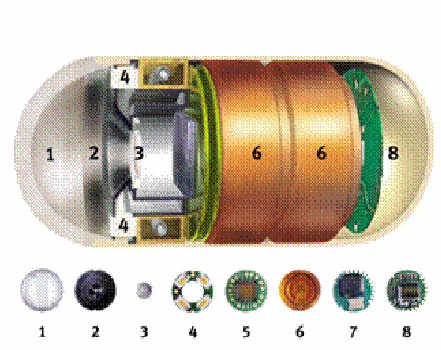


Fig.2. Components diagram of a WCE 1-Optical dome, 2-Lens holder, 3-Lens, 4-Illunimating LEDs, 5-CMOS imager, 6-Battery, 7-Transmitter, 8-Antenna.

Since WCE has been gradually applied in hospitals, some studies have been investigated towards aiding in doctors to diagnose diseases in GI tract. The authors in [3] proposed a method using color and texture to select the informative frames in WCE video. Bleeding region detection and ulcer region detection for WCE images have been investigated in [4] and [5], respectively, and the preliminary experimental results show that their proposed schemes work fine for WCE images.

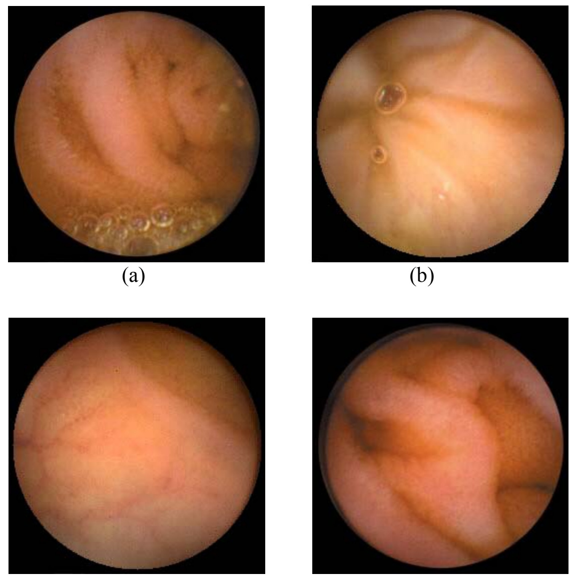
As intestinal polyp is one of the most common diseases in GI tract, we focus on intestinal polyp region recognition in this work. To achieve this goal, a comparative study scheme that exploits MPEG-7 shape descriptor and Zernike moments is carried on in this paper. Both features are invariant to rotation, translation and scale change by considering the specific imaging circumstances for WCE in the digestive tract such as the motion of the camera and so on. Experimental results on our present data show that these two schemes achieve a satisfying performance of polyp detection when using multi-layer perceptron (MLP) neural network as the classifier.

The remainder of this paper is organized as follows. The shape feature analysis using MEPG-7 region based shape descriptor and Zernike moments is discussed in the following section. Section III gives the experimental results, and we draw some conclusions and make some discussions at the end of this paper.

II. Shape Feature Analysis

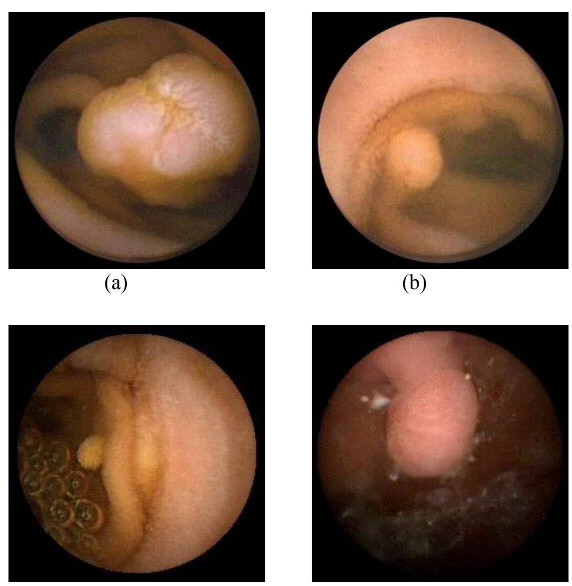
A polyp is an abnormal growth of tissue protruding from the mucous membrane. Polyps are commonly found in the intestine, stomach and so on. Intestinal polyps grow out of the lining of the small and large bowels, and they come in a variety of shapes such as round, droplet, and so on. Fig.3 shows some typical normal WCE image, and Fig.4 shows some WCE images with different polyps, which vary in shape. In fact, physicians can directly use shape features as a powerful clue to recognize polyps in images, which demonstrates that shape carries semantic information, so we concentrate on shape feature analysis for polyp detection in WCE images in our work.

In the imaging process of WCE, the images usually suffer from shape distortion due to the specific imaging circumstances such as motion of camera and mucosa, Moreover, different images from different patients may be obtained under different imaging conditions with a great deal variation in lighting and so on. Therefore, it is necessary to consider geometric deformation effects for WCE images because polyps are different when objects are viewed under different angles and different distances. A classical approach to overcome this problem is to design rotation, translation and scale invariant features. Considering this specific requirement, we turn our attention to MEPG-7 region based shape descriptor and Zernike moments.



(a) (b) (c) (d)

Fig.3 Some representative normal intestinal WCE images



(a) (b) (c) (d)

Fig.4 Some representative intestinal WCE images with polyp

A. Angular Radial Transformation-Region based Shape Descriptor

MPEG-7 visual descriptor provides a powerful tool to quantify the content in a video or an image. MPEG-7 supports two kinds of descriptors to describe the object shape, region based and contour based. As demonstrated in Fig.4, it may be very challenging to get the accurate contour of the polyps in those images in fact. As such, we resort to region-based shape descriptor.

A region based shape descriptor gives a compact and efficient approach to express the pixel distribution in a 2-D object region; and MPEG-7 adopts angular radial transform (ART) as the region based descriptor. The ART is a complex orthogonal unitary transform defined over a unit disk, which is built on complex orthogonal sinusoidal basis functions in polar coordinates. And it belongs to the broad class of shape analysis schemes built from moments [6]. For each shape, the ART coefficients F_{nm} can be obtained according to the following formula [7]:

$$F_{nm} = \int_0^{2\pi}\int_0^1 V_{nm}(\rho,\theta) f(\rho,\theta)\rho d\rho d\theta \quad (1)$$

where $f(\rho,\theta)$ is an image function represented in polar coordinates and $V_{nm}(\rho,\theta)$ is the ART basis function which is separable along the angular and radial directions:

$$V_{nm}(\rho,\theta) = \frac{1}{2\pi}\exp(jm\theta)R_n(\rho) \quad (2)$$

where $R_n(\rho) = \begin{cases} 1, & n=0 \\ 2\cos(\pi n\rho) & n \neq 0 \end{cases}$.

The ART descriptor is defined as a set of normalized magnitudes of the set of ART coefficients. Rotational invariance is obtained by using only the magnitude of these coefficients [7]. In MPEG-7, three radial functions and twelve angular functions are employed ($n < 3, m < 12$) [8]. Concerning scale normalization, the ART coefficients are divided by the magnitude of the ART coefficient with $n = 0, m = 0$.

To use ART to describe a shape in a color image, the ART descriptor can be described on the corresponding gray image, and the function $f(\rho,\theta)$ should also be normalized to the interval $[0,\ 1]$ [8].

B. Zernike moments

Taking into account the specific imaging circumstances of WCE images mentioned before, we need a shape feature that is invariant to rotation, scale and translation. Fortunately, Zernike moments also satisfy these properties. A basis function for the Zernike moment is defined by [9]:

$$V_{nm}(x,y) = V_{nm}(\rho\cos\theta, \rho\sin\theta) = R_{nm}(\rho)\exp(jm\theta) \quad (3)$$

where

$$R_{nm}(\rho) = \sum_{s=0}^{n-\frac{|m|}{2}} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)} \rho^{n-2s} \quad (4)$$

where ρ is the radius from (x, y) to the shape centroid, θ is the angle between ρ and the $x-axis$. n, m are integers that satisfy the condition $n-|m| = even,\ |m| \le n$.

Zernike polynomials are a complete set of complex-valued function orthogonals defined over the unit disk. A Zernike moment can be then defined as:

$$A_{nm}(\rho) = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V_{nm}^*(x,y) \quad x^2+y^2 \le 1 \quad (5)$$

where * represents complex conjugate. A_{nm} is a complex number. The unit disk can be centered on the center mass of the image, which makes the moments both scale and translation invariant [10]. Rotation invariance is achieved by using only the magnitude of the moments. Zernike moment raised a lot of attention as a powerful tool to describe the shape features [11][12].

Zernike moment invariant can be constructed to arbitrary order. In our implementation, we obtain up to order 10 of Zernike moment for the corresponding gray image to describe the shape feature for each WCE image, resulting in 36 feature elements in the shape feature vector.

The performance of the above two shape features for polyp detection in WCE images will be investigated through experiments, which are described in the following section.

III. EXPERIMENT RESULTS

After getting the shape feature representation for WCE images, we still need a classifier to show their discrimination ability. Since the emphasis of our present work in this paper is to investigate different shape features suitable for the purpose of computer aided detection, we just choose neural network as the classification tool. And we do not study some other classifiers especially support vector machine (SVM), which is generally considered to have a good classification performance.

Neural network has been widely used in medical decision applications because it is believed that they have great predictive ability. In this paper, we will employ the traditional multi-layer perceptron (MLP) neural network because MLP neural network has many advantages over other classifiers [13] such as better generalization ability, robust performance, and less training data. As MLP neural network is familiar to most colleagues, the introduction of MLP neural network is omitted in this paper. Details about the MLP neural network can be found in [13].

Experts of GI tract selected a data set composed of 300 representative abnormal (150) and normal (150) images from 2 patients' video data. The original images are manually labeled to provide the ground truth. The image containing any polyp region is labeled as a positive sample; otherwise, it is labeled as a negative sample. In order to prevent over-fitting of the classification results, we used three-fold cross-validation for all our classification experiments. A three-layer MLP neural network with two nonlinear outputs was employed for the purpose of classification in our experiments, the number of input nodes of the MLP depends on the number of input features. For ART shape descriptor, it is 36; and for 10 order Zernike moment, it happens to be also 36. The number of epochs for training the MLP neural network is set to 5000. The activation function in hidden layer and output layer in our experiments is the hyperbolic tangent sigmoid and the liner transfer function, respectively.

Because the number of hidden nodes has a strong impact on the final classification results, several variation on the number of hidden layer's neuron, ranging from 5 to 50 with 5 nodes' increment each time, were carried out in our experiments. Due to the space constraint, we do not show the performance of the classification result under each set of hidden neural network. Instead we only demonstrate the final average performance of the classification.

The success of classification of polyp WCE image using MLP neural network is measured by accuracy, specificity, and sensitivity, respectively, which are widely employed by the colleagues in the filed of medical decision making to assess the performance of classification. Here is the definition of them:

$$\text{Accuracy}=\frac{Number\ of\ Correct\ \ predictions}{Number\ of\ Positives+Number\ of\ Negatives} \quad (6)$$

$$\text{Specificity}=\frac{Number\ of\ Correct\ \ Negative\ \ predictions}{Number\ of\ Negatives} \quad (7)$$

$$\text{Sensitivity}=\frac{Number\ of\ Correct\ Positive\ \ predictions}{Number\ of\ Positives} \quad (8)$$

We recorded the average performance of the three-fold cross-validation experimental results in Table I. From this table, we may conclude that it is promising to employ shape features to built computer based detection system for intestinal polyp detection in WCE images. Moreover, the accuracy of the proposed scheme using Zernike moments as the shape features reaches 86.1%, together with a specificity of 82.5% and a sensitivity of 89.8%. This illustrates that Zernike moments based shape features has a better performance for polyp recognition in WCE images. The reason why Zernike moments show better discrimination ability for polyp detection may be due to the following two reasons. First, Zernike is more robust to different shape distortions since it is invariant to translation, rotation and scaling, while ART is only robust to rotation change. Moreover, Zernike moments have less information redundancy compared to ART descriptor [6].

TABLE I CLASSIFICATION RESULTS USING ART AND ZERNIKE MOMENTS(%)

	ART	ZERNIKE
Sensitivity	84.1	89.8
Specificity	72.7	82.5
Accuracy	78.4	86.1

IV. CONCLUSION

A comparative study of using two different shape features i.e., ART descriptor and Zernike moments, to detect intestinal polyps for WCE image has been conducted in this paper. As far as we know, the shape feature, which is a higher level feature than color and texture, has never been considered in the filed of capsule endoscopy images. Since the WCE images suffer from motion and so on, it is necessary to consider the effect of rotation, translation and so on. ART descriptor and Zernike moments used in this paper share one point in common, i.e., they are both invariant to rotation. Experiments on our present WCE images show that these two methods work fine for the purpose of computed aided polyp detection, and Zernike moments based features shows a better performance of detection.

Future research lies in the following three directions. First, the conjunction of shape with color will be investigated since polyps in WCE images also demonstrate different color patterns, as shown in Fig.4. In addition, more patients' data will be collected in the near future in order to testify the robustness of the aforementioned features. Furthermore, advanced classifier such as SVM can be integrated in the proposed scheme so as to further improve the classification performances of these two features.

REFERENCES

[1] S.Parker, T. Tong, S.Bolden, and P.Wingo, "Cancer Statistics 1997," CA Cancer J. Clinicians, vol.47, pp.5-27,1997.

[2] D. G. Adeler, C. J.Gostout, "Wireless Capsule Endoscopy," Hospital Physician , pp.14-22, May, 2003.

[3] M.K. Bashar, K.Mori, Y. Suenaga, T.Kitasaka, Y. Mekada, "Detecting informative frames from wireless capsule endoscopic video using color and texture features," 11th Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008, LNCS, vol.5432, pp.603-610,2008.

[4] B. Li, Max Q.-H. Meng, "Computer aided detection of bleeding regions in capsule endoscopy images," IEEE Transactions on Biomedical Engineering, Vol. 56, No. 4, April, 2009, pp.1032-1039;

[5] B. Li, Max Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," Image and Vision Computing, Vol. 27, No. 9, August, 2009, pp.1336-1342;

[6] C.H. Teh, R.T. Chin, "On Image Analysis by the Methods of Moments," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 4,, pp.496-513

[7] W.-Y. Kim,Y.-S. Kim, "A New Region-Based Shape Descriptor," ISO/IEC MPEG99/M5472, Maui, Hawaii, Dec. 1999.

[8] S.Jeannin, Mpeg-7 Visual part of experimentation Model Version 9.0. In: ISO/IEC JTC1/SC29/WG11/N3914, 55th Mpeg Meeting, Pisa, Italy. 2001

[9] M.R. Teague, "Image analysis via the general theory of Moments," J Opt Soc Am, vol.70, no.8, 1980, pp.920-930

[10] D. Zhang, G. Lu, "Evaluation of MPEG-7 shape descriptors against other shape descriptors," Multimedia Systems, vol. 9, 2003, pp.15-30.

[11] Y. Xin; M. Pawlak, S.Liao, "Accurate Computation of Zernike Moments in Polar Coordinates," IEEE Trans. Image Processing, vol.16, no.2, 2007, pp. 581 – 587

[12] J. Revaud, G. Lavoue, A. Baskurt, "Improving Zernike Moments Comparison for Optimal Similarity and Rotation Angle Retrieval," IEEE Trans. PAMI vol.31.no.4, 2009, pp. 627 – 636

[13] S.Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed, Prentice Hall, New Jersey, 1996.