World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:4, No:2, 2010

Motion Analysis for Duplicate Frame Removal in Wireless Capsule Endoscope Video

Min Kook Choi, Hyun Gyu Lee, Ryan You, Byeong-Seok Shin and Sang-Chul Lee
Department of Computer Science & Engineering, Inha University.
{kaisarulius, hglee, sclee} @ inha.edu, bleularme@naver.com, bsshin@inha.ac.kr

***Abstract* -** Wireless capsule Endoscopy (WCE) has rapidly shown its wide applications in medical domain last ten years thanks to its noninvasiveness for patients and support for thorough inspection through a patient's entire digestive system including small intestine. However, one of the main barriers to efficient clinical inspection procedure is that it requires large amount of effort for clinicians to inspect huge data collected during the examination, i.e., over 55,000 frames in video. In this paper, we propose a method to compute meaningful motion changes of WCE by analyzing the obtained video frames based on regional optical flow estimations. The computed motion vectors are used to remove duplicate video frames caused by WCE's imaging nature, such as repetitive forward-backward motions from peristaltic movements. The motion vectors are derived by calculating directional component vectors in four local regions. Our experiments are performed on small intestine area, which is of main interest to clinical experts when using WCEs, and our experimental results show significant frame reductions comparing with a simple frame-to-frame similarity-based image reduction method.

Keywords – Wireless capsule endoscopy, optical flow, duplicated image, duplicated frame.

I. INTRODUCTION

WIRELESS Capsule Endoscopy (WCE) has been considered as the most convenience method to perform medical diagnosis on many diseases in small bowels (see Figure 1). However, one of the main obstacles for efficient diagnosis is that it requires huge amount of effort due to long inspection time for clinical experts because of the WCE's imaging nature which relies solely on peristalsis [1]. For example, during approximately 8-hour examination, over 55,000 images are taken and transmitted to outer device for later examination. Typically, it is known that even a highly trained clinical expert spends more than two hour to inspect the entire sequence of a WCE video [2]. Therefore, computer aided inspection system supporting frame summarization based on intelligent removal of duplicated frames is urgently needed.

Recently, many researchers have put large amount of effort on analyzing WCE images to support computer aided diagnosis system. Hwang took probabilistic approach for extracting color feature to detect blood showing about 92% detection ratio [2]. Others used texture and chromaticity histogram to find blood and ulcer from WCE images, and achieved about 60-85% detection ratio [4, 5]. Beside from simple blood detection, some researchers proposed a method to discriminate different tissues in a human digestive system using Discrete Cosine Transform (DST) and Principal Component Analysis (PCA) [3]. Eventually, the main goals of the existing research are in automated detection of some interesting medical features in clinical diagnosis.

In our research, instead of tackling to automated detection problem, which is somewhat limited to several applications, we address a more general and widely applicable problem, which efficiently visualizes the large data which could ultimately help an expert to inspect the data efficiently. In particular, our goal is to reduce the amount of data for inspection while minimizing the information loss. Specifically, the support for efficient computer-aided visualization system could be approached by data reduction through duplicate frame removal taken in duplicate spatial location in the subject's digestive system. In this paper, we propose a strategy to reduce such duplicated video frame by analyzing motion vectors from WCE. Our experimental results show significant data reduction over a method based on simple frame-to-frame similarity.

In the following section, we describe the optical flow estimation method, and then show an energy based formulation [9]. In section 3, we show a scheme for analyzing flow vectors. An experimental result is presented in Section 4, followed by a conclusion in Section 5.

Figure 1 Wireless Capsule Endoscope (WCE) and an example image taken

II. OPTICAL FLOW ESTIMATION

In this paper, we adopted an optical flow estimation algorithm based on [6]. For optical flow estimation many researchers typically used gray value constancy assumption, which is, the gray value of a pixel is not changed by the displacement, such as:

$$I(x, y, t) = I(x + u, y + v, t + 1) \quad (1)$$

Nevertheless, the grey value constancy assumption has some defect that is quite susceptible to slight change of intensity in brightness. The optical flow estimation is further improved as the way that the result is more robust against grey value changes [10] as follows:

$$\nabla I(x, y, t) = \nabla I(x + u, y + v, t + 1) \quad (2)$$

Although the constraint (2) is particularly helpful for

translator motion referring to gradient constancy assumption, the abovementioned two constraints are too local, only considering motions of the order of one pixel. Therefore, it is necessary to consider neighbor pixels to detect optical flows in large displacements considering smoothing and multi-scale approach. These model assumptions are achieved by energy function. The main energy functional is expressed as:

$$E(u,v) = E_{Data}(u,v) + \propto E_{Smooth}(u,v) \quad (3)$$

$$E_{Data}(u,v) = \int_{\Omega} \Psi(|I(x+w) - I(x)|^2 + \gamma|\nabla I(x+w) - \nabla I(x)|^2)dx \quad (4)$$

$$E_{Smooth}(u,v) = \int_{\Omega} \Psi(|\nabla_3 u|^2 + \gamma|\nabla_3 v|^2)dx \quad (5)$$

$\propto$ represents the regularization term and γ is weight term determining the importance of gradient value in the image. An increasing concave function $\Psi(s^2)$ is applied for robust energy to prevent that outliers get too much influence on the estimation [11, 12]. Equation (5) represent the model assumption of a piecewise smooth flow field [13, 14].

We applied minimization of the energy function as introduced in Horn and Schunck [6]. It minimizes the above functional globally as follows.

$$I_z^{k+1} \approx I_z^k + I_x^k du^k + I_z^k dv^k \quad (6)$$

However, equation (6) is still highly non-linear due to the terms of (du^k, dv^k), Euler-Lagrange equation is used to linearize the equation (3) followed by applying a numerical approximation such as fixed point iteration [9]. The linearization of the functional (3) yields

$$0 = (\Psi')^{k,l}_{Data} \cdot \left(I_x^k\left(I_z^k + I_x^k du^{k,l+1} + I_y^k dv^{k,l+1}\right) + \gamma I_{xx}^k\left(I_{xx}^k + I_{xx}^k du^{k,l+1} + I_{xy}^k dv^{k,l+1}\right) + \gamma I_{xy}^k\left(I_{yz}^k + I_{xy}^k du^{k,l+1} + I_{yy}^k dv^{k,l+1}\right)\right) - \propto \mathrm{div}((\Psi')^{k,l}_{Smooth}\nabla_3(u^k + du^{k,l+1})) \quad (7)$$

The resulting linear system of equation can be solved by numerical methods such as Gauss-Seidel or SOR (successive over-relaxation method) iterations. These schemes are usually achieved by coarse-to-fine warping techniques [7, 8, 11]. In [9], it is proved that the warping technique implements the minimization problem of a non-linearized constancy assumption mathematically.

III. VECTOR FLOW ANALYSIS

To estimate movement of WCE, we calculate the set of flow vectors in particular region. We assume that WCE movements, such as forward, backward and spin, differently influence the scene depending on region. Figure 2 shows the flow vectors in the considered regions between an adjacent frames in particular region. Each flow vectors are defined as Euclidian distance of x, y components, $u = \sqrt{(u_x{}^2 + u_y{}^2)}$.

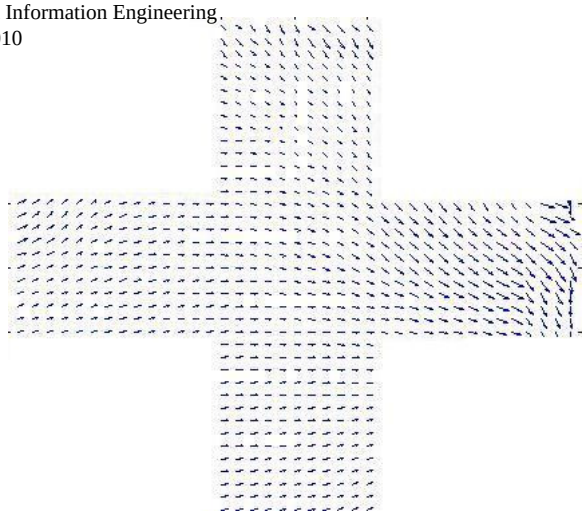

Figure 2 Estimated flow depend on region

To analyze motion of WCE, we combine vector in each region by averaging all flow vectors over all the image sequences. The top (north) region is calculated as the equation (8):

$$U_R = \sqrt{\left(\frac{1}{N_R}\sum_0^{N_R} u_x\right)^2 + \left(\frac{1}{N_R}\sum_0^{N_R} u_y\right)^2} \quad (8)$$

where R = {top, bottom, left, right}. These average values are used to determine the threshold to filter out unreasonable data, i.e., small optical flow value. Based on our experimental observations, we set a non-change movement threshold to the value of under 0.01 and outlier over 10.0. It showed that these thresholds effectively reduced the frames estimating non-moving (duplicated) images.

Next, to estimate the overall directional tendency of the optical flows in each region, we considered the angular vectors calculated by all of arctangent values with x, y component in each region.

$$Ang_R = \tan^{-1}\left(\frac{\sum_0^{N_R} u_x}{\sum_0^{N_R} u_y}\right) \quad (9)$$

where R = {top, bottom, left, right}.

The extracted angular vectors are further quantized into eight regions in 45 degrees at range such that 337.5° ~ 22.5° is mapped to R, 22.5° ~ 67.5° to RT, 67.5° ~ 112.5° to T, 112.5° ~ 157.5° to LT, 157.5° ~ 202.5° to L, 202.5° ~ 247.5° to LD, 247.5° ~ 292.5° to D, 292.5° ~ 337.5° to RD (see Figure 3). We call these quantized regions as *directional components*.

By combining the directional components, we derived the combinations of movement patterns of WCE. Although we could consider all combinations (16777216(8^8) kinds) of the directional components in four regions, it is not practical to apply for our application. Figure 5 shows the histograms of the directional components in each region. In each region, we extract the major directional component with highest population for each movement categories (i.e., forward, backward, left, right, up, down, clockwise and counter clockwise). From the Figure 5, we build a set of

directional component vectors for each region (v_1,v_2,v_3,v_4) that fall into each movement categories (forward, backward, etc.). The vectors are constructed by taking the maximum directional component (T, RT, R, etc.) in each region.

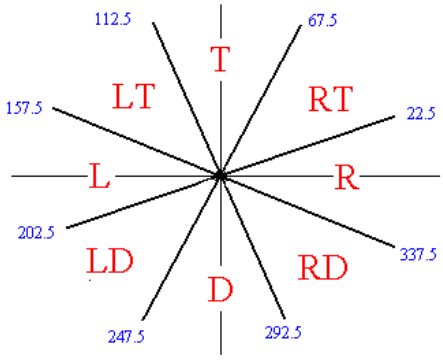

Figure 3 Range of angles on x, y coordinate

Among the set of directional component vectors, as mentioned earlier, we have selected a subset of the directional component vectors consisting of 16 meaningful vectors which showed relatively accurate WCE movement estimation as shown in Table 1. In Table 1, we choose four useful movement categories for our application purpose (Forward, Backward, Clockwise, and Counter Clockwise). Each pattern has been tested 3 times using our simulation results performed to a synthetic data shown in Figure 4 (a) and (b).

Table 1 shows the result of classification of the WCE movement based on the directional flow vector. The first column represents the classification categories, the second column the selected directional flow vectors and third column the proportional factor λ_p of the flow vectors, respectively. We defined λ_p as a proportional factor of flow vector, as $\lambda_p = v(h)/s(h)$, where $v(h)$ is inlier vectors and $s(h)$ is the total vectors in the regional vector population.

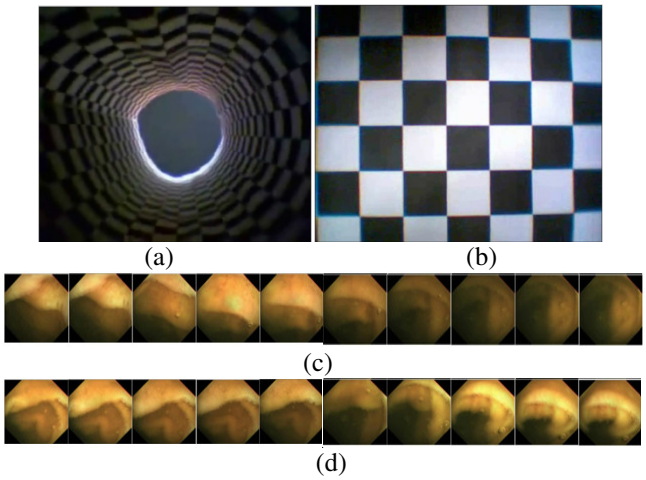

Figure 4 (a) synthetic image for extracting forward and backward movement pattern and (b) for parallel and rotational movement. (c) and (d) are WCE image moving forward and backward.

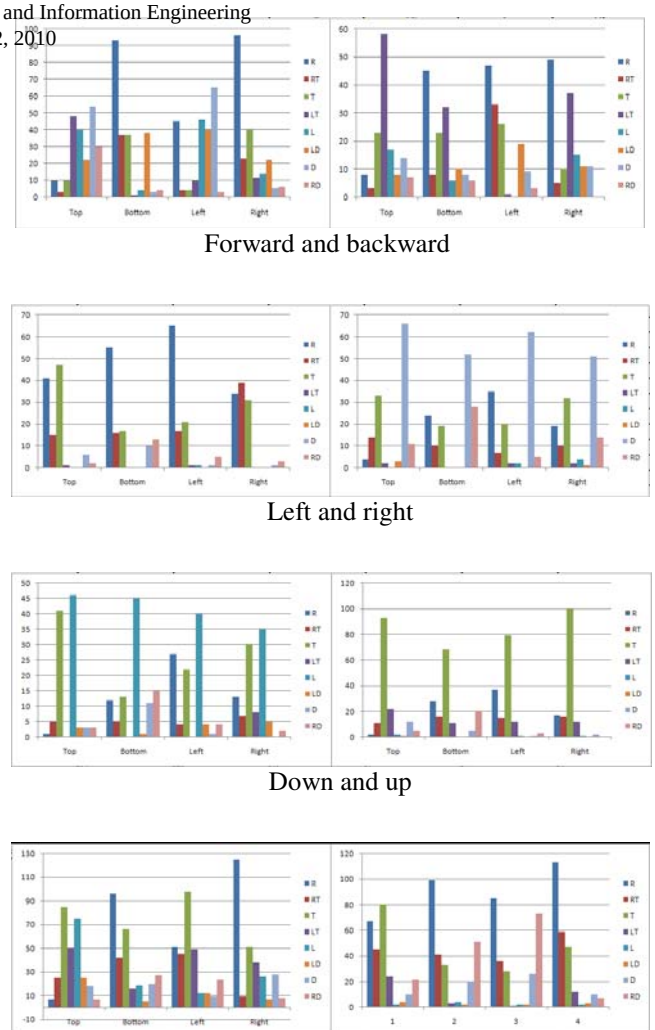

Figure 5 Histogram result from testing image sequences. There are 8 kinds of movement patterns. x axis shows Top, Bottom, Left and Right regions of angle flows. y axis shows the histogram results. Each parameter has the movement pattern, R, RT, T, LT, L, LD, D and RD from left.

	Directional flow vector	λ_p
Forward	(D, R, D, R)	35.48
	(LT, R, L, R)	32.60
Backward	(LT, R, R, R)	35.69
	(LT, LT, RT, LT)	28.99
Clockwise	(T, R, T, R)	34.59
	(L, R, T, R)	33.37
Counter Clockwise	(T, R, R, R)	37.25
	(R, RD, RD, R)	30.04

Table 1 Categorization of movement pattern using histogram data of figure 5. We choose 8 kinds of significant movement patterns.

Since the main goal of our work is in removing the duplicate image frames by detecting the backward motion of the WCE, we further focused on the backward movement. Figure 6 and Table 2 show the histogram and the table re computed for the backward movement only. 4.

IV. EXPERIMENT

We performed an experiment with real WCE images about 10,000 frames recoded in small intestine. We assume

that backward movement contains frames taken in duplicated location in intestine. After removal of such frames in sequences, they are recorded for visualization.

Figure 6 shows the frames regarded as backward or non-change movement, and nearby 3 frames. Through the all sequence 10,000 images, we detected 74 frames containing backward motions. Considering the entire small intestine, in 30,000 ~ 50,000 frames, the detected backward motions are about 222 ~ 370 frames. So data reduction of WCE images can achieved about 2200~ 4000 frames in total intestine (7.4 ~ 8%). In our experiment, successful detection of duplicated recorded location images are 65 frames, it is about 87.84%.

Although they do not detect all the backward motion frames, we observed that the nearby frames of detected image are usually the continuation of the backward movements with small flow vector magnetite. It is implied that the method can be dramatically improved further using the detected frame as a key frame for nearby frame clustering. Figure 7 shows the example of detected frames.

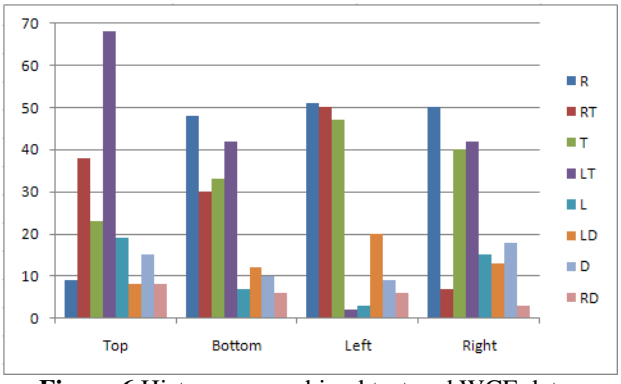

Figure 6 Histogram combined test and WCE data.

	Directional flow vector	λ_p
Backward	(LT, R, R, R)	28.86
	(LT, LT, LT, LT)	26.86
	(LT, LT, T, T)	26.20
	(LT, T, T, LT)	25.27
	(RT, RT, RT, T)	20.61

Table 2 Categories of backward movement from test and WCE image data.

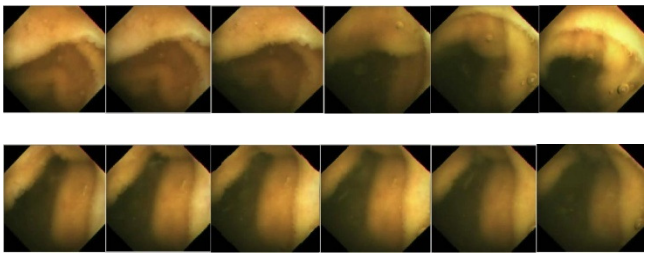

Figure 7 Detected image and nearby 3 frames.

V. CONCLUSION

We proposed a method to remove duplicated video frames in WCE using optical flow estimation. Especially we focused on the backward movement of the WCE which contains large amount of duplicate frames. The proposed method computed the directional component vectors by analyzing the population of the optical flows in four pre-defined regions. Experiment could be further performed by some improvement such as re-defining range of angles and flow region in vector as well as magnitude threshold. This idea can also be widely used to detect motions in other domains producing a video containing repetitive video frames such as vehicle mounted cam, surveillance movie, real and real-time navigation system, etc.

REFERENCES

[1] G. Iddan, G. Meron, A. Glukhovsky, P. Swain, "Wireless capsule endoscopy", *Nature*, vol. 405, Issue 6785, pp. 417-418, 2000.
[2] 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*, vol. 6144, pp. 577-587, 2006.
[3] J. Berens, M. Mackiewicz, D. Bell, "Stomach, intestine and colon tissue discriminators for wireless capsule endoscopy images", Proceedings of SPIE, Conference on Medical Imaging, vol. 5747, pp. 283-290, 2005.
[4] B. Li, MQH. Meng, "Texture analysis for ulcer detection in capsule endoscopy images", *Image and Vision Computing,* vol. 27, pp. 1336-1342, 2009.
[5] B. Li, MQH. Meng, "Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments", *Computer in Biology and Medicine*, pp. 141-147, 2009.
[6] Berthold K. P. Horn, Brian G. Schunck, "Determining Optical Flow", *Artificial Intelligence*, pp. 185-203, 1981.
[7] P. Anandan, "A computational framework and an algorithm for the measurement of visual motion", *International Journal of Computer Vision*, 2, pp. 283-310, 1989.
[8] E. Memin, P. Perez, "Hierarchical estimation and segmentation of dense motion fields", *International Journal of Computer Vision*, 46(2), pp. 129-155, 2002.
[9] T. Brox, A. Bruhn, N. Papenberg, J. Weickert, "High accuracy optical flow estimation based on a theory for warping", *European Conference on Computer Vision,* LNCS 3024, pp. 25-36, 2004.
[10] S. Uras, F. Girosi, A. Verri, and V. Torre. "A computational approach to motion perception", *Biological Cybernetics*, 60, pp. 79-87, 1988.
[11] M. J.Black. P. Anandan, "The robust estimation of multiple motions: parametric and piecewise smooth flow fields", *Computer Vision and Image Understanding,* 63(1), pp. 75-104, 1996.
[12] E. Memin. P. Perez, "A multigrid approach for hierarchical motion estimation", *In Proc. Sixth International Conference on Computer Vision*, pp. 933-938, 1998
[13] L. I. Rudin, S. Osher, E. Fatemi, "Nonlinear total variation based noise removal algorithms", *Physica D*, 60, pp. 259-268, 1992.
[14] I. Cohen, "Nonlinear variational method for optical flow computation", *Proc. Eighth Scan-dinavian Conference on Image Analysis*, volume 1, pp. 523-530, 1993
[15] L. Alvarez, J. Esclarin, M. Lefebure, J. Sanchez, "A PDE model for computing the optical flow", In *Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicationes*, pp.1349-1356, 1999., 1999.