35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Application of the Real-Time Retinex Image Enhancement for Endoscopic Images

Hiroyuki Okuhata, *Member, IEEE,* Hajime Nakamura, Shinsuke Hara, *Member, IEEE*, Hiroshi Tsutsui, *Member, IEEE,* and Takao Onoye, *Member, IEEE*

Abstract— **This paper presents a real-time image enhancement technique for gastric endoscopy, which is based on the variational approach of the Retinex theory. In order to efficiently reduce the computational cost required for image enhancement, processing layers and repeat counts of iterations are determined in accordance with software evaluation result, and as for processing architecture, the pipelining architecture can handle high resolution pictures in real-time. To show its potential, performance comparison between with and without the proposed image enhancement technique is shown using several video images obtained by endoscopy for different parts of digestive organ.**

I. Introduction

Endoscopy means looking inside the body for medical purposes using a compact device [1]. An endoscope is equipped with a CCD or CMOS image sensor chip and a flexible tube containing a light source at its tip, and can take a sequence of digital pictures and sometimes has a capability of zooming-in at an interesting region for examining its smaller and finer details. It is really a power tool for internal diagnosis and inspection, but using an endoscope is relatively uncomfortable for a patient, introducing the possibility of perforation of organs, infection and hemorrhage. Endoscopy is sometimes limited for the inspection on the esophagus to the stomach; it is not applied for the small intestine, because the small intestine is very long and convoluted. On the other hand, a wireless capsule endoscope, which is equipped with a light source, lens, camera, and radio transmitter, is a powerful inspection tool for the small intestine, especially for detecting the cause of gastrointestinal bleeding. It automatically takes a number of pictures during traveling through the digestive systems, just propelled by peristalsis.

Experienced physicians and technicians have been responsible for their diagnosis based on available information provided by wired and wireless endoscopy, and as the number of images to be inspected increases, the role of such medical specialists becomes the most time consuming and exhausting. For instance, a wireless capsule endoscope takes two pictures per second in two hours of traveling time through the small intestine, so it results in more than 50,000 images. Therefore, computer-aided support systems are essential, which automatically detect relevant images and annotate them for further analysis by a physician. However, the quality of endoscopic images is too low to put directly to such systems, so they need some pre-processing.

Endoscopic images suffer from various kinds of degradations. One major cause is undesired noise such as thermal noise in CCD or CMOS chips. To reduce the noise, several efficient filtering techniques have been proposed. Another major cause is inhomogeneous brightness and poor contrast, and related to this, specular highlights are a problem inherent to wireless capsule endoscope, which comes from its nature that it cannot steer the orientation of the camera. Furthermore, gastric organ has its own convoluting, bending and waving natures, so inhomogeneous brightness and poor contrast inevitably occur in its images.

To overcome the inhomogeneous brightness and poor contrast and to reduce highlighted areas as much as possible, gamma correction, masking and histogram equalization, which are categorized into an image enhancement technique, have been used for long years. They are implemented in simpler algorithms, however, they sometime do not work well, because they just try to compensate for each pixel value uniformly based on given equations. On the other hand, an adaptive image enhancement technique always works well because it refers surrounding pixels to reproduce a high quality image, however, it suffers from high computational cost. Therefore, effective reduction of the computation cost is essential for the practical use of adaptive image enhancement technique in wired and wireless endoscopy.

Among the adaptive image enhancement schemes, the Retinex theory, which introduces a model of the lightness and color perception of human visual system, is well known as a powerful tool applicable for still and moving images. So far, in the Retinex theory, we have applied a variational model algorithm for enabling iterative processing and a pipelining architecture for realizing real-time processing. In this paper, we outline the Retinex theory image enhancement technique and discuss its potential in the application for endoscopic images.

II. Retinex Theory and Variational Model

A. Theory

The Retinex theory, which is based on a property of the color constancy phenomenon, decomposes a given image into two different images: the reflectance image and the illumination image. Here, color constancy phenomenon

Manuscript received on January 10, 2013.

H. Okuhata is with Synthesis Corporation, 2-6-9, Awajimachi, Chuo-ku, Osaka 541-0047, Japan (phone: +81-6-6223-5841; fax: +81-6-6223-5844; e-mail: okuhata@synthesis.co.jp).

H. Nakamura is with the graduate school of medicine, Osaka City University, 1-5-17 Asahimachi, Abeno-ku, Osaka 545-0051, Japan.

S. Hara is with the graduate school of engineering, Osaka City University, 3-3-138, Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan.

H. Tsutsui is with the graduate school of informatics, Kyoto University, Yoshida-Honmachi, Sakyo, Kyoto 606-8501, Japan.

T. Onoye is with the graduate school of information science and technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan.

stands for the fact that human visual system can practically recognize and match colors under a wide range of different illuminations. The Retinex theory [2-8] utilizes this property to extract the illumination image. In the Retinex theory, an input image $S(x,y)$ can be expressed by the following equation;

$$S(x, y) = L(x, y) \cdot R(x, y) \ , \qquad (1)$$

where $L(x,y)$ and $R(x,y)$ indicate the illumination image and the reflectance image, respectively. Therefore, image enhancement can be achieved by extracting $L(x,y)$ from $S(x,y)$ in order to generate $R(x,y)$ as the illumination -independent image.

The original Retinex theory does not explicitly specify a scheme to estimate the illumination image, which is in general an ill-posed problem. Several illumination models have been so far proposed for the Retinex theory [2-8]. Among them, we focus on the variational model as the most viable for practical applications in terms of computational cost and image quality.

B. Variational Model

Variational model algorithm is constructed to minimize the following cost function;

$$F[l] = \int_{\Omega} \left(\left|\nabla l\right|^2 + \alpha (s - l)^2 + \beta \left|\nabla (s - l)\right|^2 \right) dxdy \qquad (2)$$

where, α and β are parameters, $s(x,y)$ and $l(x,y)$ represent the logarithmic expression of input image, reflectance image, and illumination image, respectively. Here, the logarithmic expression of reflectance image can be expressed by $r(x,y)=s(x,y)-l(x,y)$. Penalty terms, such as $|\nabla l|^2$, $(s-l)^2$, and $|\nabla (s-l)|^2$ represent spatial smoothness of the illumination image, closeness between l and s, and spatial smoothness of the reflectance image r, respectively.

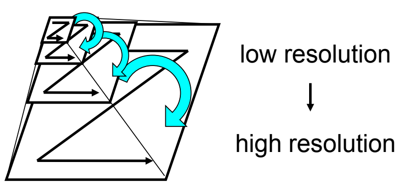


Figure 1: Layer hierarchy for illumination estimation.

To find illumination image l which minimizes the penalty $F[l]$, a projected normalized steepest descent (PNSD) algorithm is used, in which the following steps are iteratively executed from low resolution layer to high resolution layer as illustrated in Fig.1. Note that G is the gradient of $F[l]$, and the layer k represents k-times decimated image. In the following, we use images in logarithmic space, such as s, r, and l, instead of S, R, and L.

1. Input image $s(x,y)$, parameters α and β, and the number of iterations T_k for each layer k are set.
2. Gaussian pyramid of the image s, as shown in Fig. 2, is constructed by decimation with the following low-pass filter K_{PYR}. Here, k-times decimated image is denoted by $s^{(k)}$. Note that $s^{(0)}=s$.

$$\kappa_{\mathrm{PYR}} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad (3)$$

3. The following is iteratively executed for each layer k from the lowest resolution layer N. Before the following iteration, the initial illumination image in layer N, which is denoted by $l_0^{(N)}$, is initialized by $l_0^{(N)}=\max\{s^{(N)}\}$. The iteration index is denoted by j, and $j=1,2,..., T_k$.
   (a) Calculation of gradients for the illumination image and input image.

$$G_A(x, y) = \nabla l_{j-1}^{(k)}(x, y) \, , \qquad (4)$$

$$G_B(x, y) = \nabla s^{(k)}(x, y) \, , \qquad (5)$$

$$\begin{aligned} G(x, y) = & -G_A(x, y) \\ & + \alpha\left(l_{j-1}^{(k)}(x, y) - s^{(k)}(x, y)\right), \\ & + \beta\left(G_A(x, y) - G_B(x, y)\right) \end{aligned} \qquad (6)$$

where the Laplacian ∇ is defined by the following,

$$\nabla l = l * k_{\mathrm{LAP}} \, 2^{-2k} \, , \qquad (7)$$

$$k_{\mathrm{LAP}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} . \qquad (8)$$

   (b) Calculation of μ_{NSD}.

$$\mu_A = \sum_x \sum_y G(x, y)^2 \, , \qquad (9)$$

$$\mu_B = -\sum_x \sum_y G(x, y) \nabla G(x, y) \, , \qquad (10)$$

$$\mu_{\mathrm{NSD}} = \frac{\mu_A}{\alpha\mu_A + (1+\beta)\mu_A} . \qquad (11)$$

   (c) Update of the illumination image with the constraint.

$$l_j^{(k)}(x, y) = \max\left(l_{j-1}^{(k)}(x, y) - \mu_{\mathrm{NSD}} G(x, y), s^{(k)}(x, y)\right) . \qquad (12)$$

4. After T_k iterations, initialization of the next resolution layer is carried out. The initial illumination image of the next layer k-1, which is $l^{(k-1)}$, is determined by up-scaling the resulted illumination image $l_{Tk}^{(k)}$
5. The illumination image of the layer 0, which would be $l_{T0}^{(0)}$, whose resolution is the same as the input image, is finally output.

After the above mentioned steps, the reflectance image R can be obtained by

$$R = \exp(s - l) , \tag{13}$$

where s and l are the input image and the resulted illumination image, respectively. However, the reflectance image is sometimes an over-enhanced image. To avoid this over-enhancement, in [6], applying a gamma corrected illumination back to the reconstructed reflectance image. The flow of this Retinex image enhancement with illumination correction is shown in Fig.2, where the gamma correction is expressed by

$$L = L^{1/\gamma} . \tag{14}$$

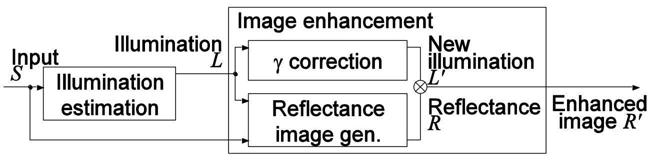


Figure 2: Flow of the Retinex image enhancement.

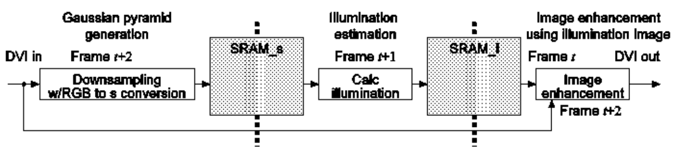


Figure 3: block diagram of the proposed architecture.

Figure 4: A picture of the image enhancement board.

III. Proposed Architecture and Its Implementation

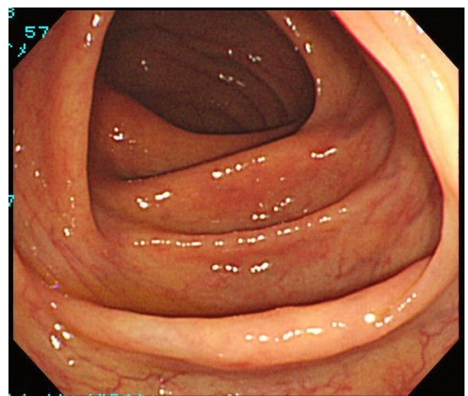
The proposed architecture and video image enhancement scheme are based on the former implementations [9,10], where N=5 and T_k={0,0,0,10,20,30}, k=0,...,5. In this scheme, it has been revealed that omission of the illumination estimation for high resolution layers contributes drastic reduction of the computational cost and hardware resources, with slight degradation of the enhancement quality.

The block diagram of the proposed architecture is shown in Fig. 3. The proposed architecture consists of three parts; (1) Gaussian pyramid generation part, (2) illumination estimation part, and (3) image enhancement part. Since these parts work in pipeline manner, each part processes its corresponding frame. In this paper, we assume that a frame buffer which stores input frames temporally is not available. Therefore, as shown in the figure, an input frame t+2 is enhanced by using the illumination estimated from a preceding frame t. When a frame buffer, which can store two successive input frames, is available, each input frame can be properly enhanced. In this case, the size of frame buffer should be 1,920×1,200×3×2=13.8 MBytes, assuming 24 bit RGB color representation [11].

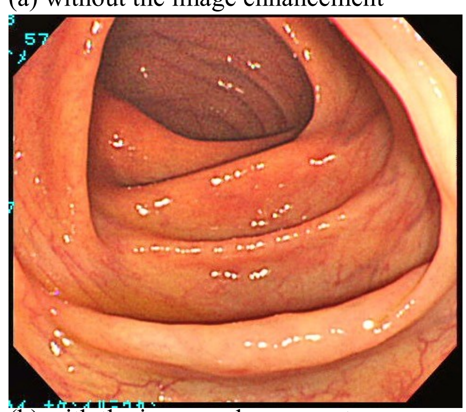
Finally, Fig. 4 shows a picture of the implemented board of the real-time Retinex image enhancement. It is developed using a Stratix III FPGA Development Kit and an FPGA device of EP3SL150F1152C2.

IV. Experiments and Discussions

The main advantage of the proposed image enhancement technique is its capability of on-line/real-time processing for video image obtained during inspection as well as off-line processing for stored video image, but in the following, using stored images by endoscopy for three different parts of digestive organ, we compare one-shot images with and without the proposed technique. The images were obtained in Osaka City University Hospital, and using the developed FPGA board shown in Fig. 4, we processed them, where the parameter of γ was optimized for each of the three images.



(a) without the image enhancement

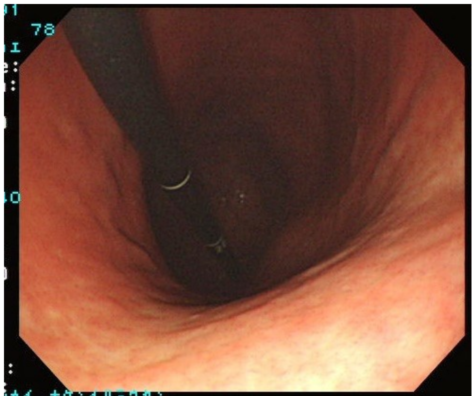


(b) with the image enhancement

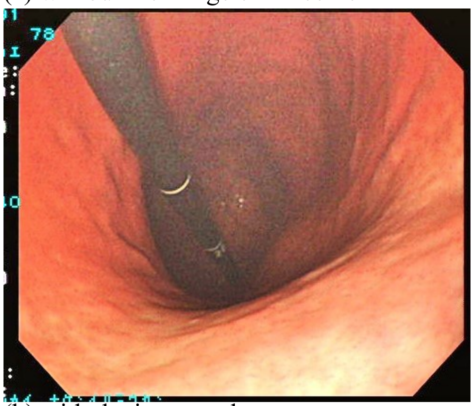
Figure 5: Image for the descending colon.

Fig. 5 compares an image for the descending colon. The raw image contains dark and bright parts due to its waving nature. Giving no significant change to the bright part, the proposed technique can successfully remove the shade only from the dark part; the proposed technique makes it possible to observe the splenic flexure in more detail.

Fig. 6 compares an image for a body of the stomach. In the raw image, we cannot distinguish the tube of the endoscope in the dark inner part, but the proposed technique can remove



(a) without the image enhancement



(b) with the image enhancement

Figure 6: Image for a body of the stomach.

(a) without the image enhancement

(b) with the image enhancement

Figure 7: Image for another body of the stomach.

the shade only from it and the shape of the tube becomes visible in the unshaded inner part of the stomach; the proposed enhancer makes it possible to observe the fornix in more detail.

Finally, Fig. 7 compares an image for another body of the stomach. In the raw image, we cannot find a polyp located at the center of the inner part, which is likely to be missed during endoscopic inspection. However, using the proposed technique, we can exactly identify it there.

V. CONCLUSION

This paper has applied the Retinex image enhancement technique to video images obtained by gastric endoscopy, where the variational model approach allows its real-time processing with low computational complexity. It has been confirmed that, by comparing images with and without the image enhancement, the proposed technique can work well for endoscopic images. It can unshade only dark parts of the images, which are unavoidable due to the waving, convoluting and bending natures of digestive organ. Therefore, the Retinex real-time image enhancement is useful and promising for endoscopy to eliminate miss of disease during long and exhausting inspection.

REFERENCES

[1] M. Liedlgruber, "Endoscopic image processing - an overview," in Proc. 6th International Symposium Image and Signal Processing and Analysis (ISPA) 2009, pp. 707-714, Sep. 2009.

[2] E. H. Land and J. J. McCann, "Lightness and Retinex theory," Journal of the Optical Society of America, vol. 61, Dec. 1964.

[3] D. J. Jobson, Z. Rahman, and A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," IEEE Transactions on Image Processing, vol. 6, no. 7, pp. 965-976, July 1997.

[4] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround Retinex," IEEE Transactions on Image Processing, vol. 6, no. 3, pp. 451-462, Mar. 1997.

[5] D. Terzopoulos, "Image analysis using multigrid relaxation methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 2, pp. 129-139, Mar. 1986.

[6] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Soble, "A variational framework for Retinex," International Journal of Computer Vision, vol. 52, no. 1, pp. 7-23, Apr. 2003.

[7] S. Carrato, "A pseudo-Retinex approach for the visualization of high dynamic range images," in Proc. 5th European Cooperation in the Field of Scientific and Technical Research (COST) 276 Workshop, pp. 1-6, Oct. 2003.

[8] C. T. Lin and C. H. Huang, "CNN-based Retinex technology," in Proc. European Conference on Circuit Theory and Design (ECCTD '03), vol. 2, pp. 69-72, Sep. 2003.

[9] K. Takahashi, Y. Nozato, H. Okuhata, and T. Onoye, "VLSI architecture for real-time Retinex video image enhancement," in Proc. Workshop on Synthesis And System Integration of Mixed Information technologies (SASIMI 2007), pp. 81-86, Oct. 2007.

[10] H. Okuhata, K. Takahashi, Y. Nozato, T. Onoye, and I. Shirakawa, "Video image enhancement scheme for high resolution consumer devices," in Proc. 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008), pp. 639-644, Mar. 2008.

[11] H. Tsutsui, H. Nakamura, R. Hashimoto, H. Okuhata, and T. Onoye, "An FPGA implementation of real-time Retinex video image enhancement," Proc. World Automation Congress (WAC) 2010. pp. 1-6, Sep. 2010.