Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2011, Article ID 257095, 15 pages
doi:10.1155/2011/257095

Research Article

A Subsample-Based Low-Power Image Compressor for Capsule Gastrointestinal Endoscopy

Meng-Chun Lin[1] and Lan-Rong Dung[2]

[1] *Department of IC Design, Avisonic Technology Corporation, No. 12, Innovation 1st Road Hsinchu Science Park, Hsinchu 300, Taiwan*
[2] *Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan*

Correspondence should be addressed to Meng-Chun Lin, asurada.ece90g@nctu.edu.tw

Received 4 August 2010; Revised 8 November 2010; Accepted 4 January 2011

Academic Editor: Dimitrios Tzovaras

In the design of capsule endoscope, the trade-offs between battery-life and video-quality is imperative. Typically, the resolution of capsule gastrointestinal (GI) image is limited for the power consumption and bandwidth of RF transmitter. Many fast compression algorithms for reducing computation load; however, they may result in a distortion of the original image, which is not suitable for the use of medical care. This paper presents a novel image compression for capsule gastrointestinal endoscopy, called GICam-II, motivated by the reddish feature of GI image. The reddish feature makes the luminance or sharpness of GI image sensitive to the red component as well as the green component. We focus on a series of mathematical statistics to systematically analyze the color sensitivity in GI images from the RGB color space domain to the two-dimensional discrete-cosine-transform spatial frequency domain. To reduce the compressed image size, GICam-II downsamples the blue component without essential loss of image detail and also subsamples the green component from the Bayer-patterned image. From experimental results, the GICam-II can significantly save the power consumption by 38.5% when compared with previous one and 98.95% when compared with JPEG compression, while the average peak signal-to-noise ratio of luminance (PSNRY) is 40.73 dB.

1. Introduction

Gastrointestinal (GI) endoscopy has been popularly applied for the diagnosis of diseases of the alimentary canal including Crohn's Disease, celiac disease, and other malabsorption disorders, benign and malignant tumors of the small intestine, vascular disorders, and medication-related small bowel injury. There are two classes of GI endoscopy: wired active endoscopy and wireless passive capsule endoscopy. The wired active endoscopy can enable efficient diagnosis based on real images and biopsy samples; however, it causes discomfort for the patients to push flexible, relatively bulky cables into the digestive tube. To relief the patients' discomfort, wireless passive capsule endoscopes are being developed worldwide [1–6].

The capsule moves passively through the internal GI tract with the aid of peristalsis and transmits images of the intestine wirelessly. Developed by Given Imaging Ltd., the PillCam capsule is a state-of-the-art commercial wireless capsule endoscope product. The PillCam capsule transmits the GI images at a resolution of 256-by-256 8-bit pixels and the frame rate of 2 frames/sec (or fps). Because of its high mobility, it has been successfully utilized to diagnose diseases of the small intestine and alleviate the discomfort and pain of patients. However, based on clinical experience; the PillCam still has some drawbacks. First, the PillCam cannot control its heading and moving direction itself. This drawback may cause image oversights and overlook a disease. Second, the resolution of demosaiced image is still low, and some interesting spots may be unintentionally omitted. Therefore, the images will be severely distorted when physicians zoom images in for detailed diagnosis. The first drawback is the nature of passive endoscopy. Some papers have presented approaches for the autonomous moving function [7–11]. Very few papers address solutions for the second drawback. Increasing resolution may alleviate the second problem; however, it will result in significant power consumption in RF transmitter. Hence, applying image compression is

necessary for saving the power dissipation of RF transmitter [12–20].

Our previous work [14] has presented an ultralow-power image compressor for wireless capsule endoscope. It helps the endoscope to deliver a compressed 512-by-512 image, while the RF transmission rate is at 1 megabit ($(256 \times 256 \times 2 \times 8)/1024^2$) per second. No any references can clearly define how much compression is allowed in capsule endoscope application. We define that the minimum compression rate is 75% according to two considerations for our capsule endoscope project. The first consideration is that the new image resolution (512-by-512) that is four times the one (256-by-256) of the PillCam can be an assistant to promote the diagnosis of diseases for doctors. The other one is that we do not significantly increase the power consumption for the RF circuit after increasing the image resolution from the sensor. Instead of applying state-of-the-art video compression techniques, we proposed a simplified image compression algorithm, called GICam, in which the memory size and computational load can be significantly reduced. The experimental results show that the GICam image compressor only costs 31 K gates at 2 frames per second, consumes 14.92 mW, and reduces the image size by at least 75%.

In applications of capsule endoscopy, it is imperative to consider the tradeoffs between battery life and performance. To further extend the battery life of a capsule endoscope, we herein present a subsample-based GICam image compressor, called GICam-II. The proposed compression technique is motivated by the reddish feature of GI image. We have previously proposed the GICam-II image compressor in paper [21]. However, the color importance of primary colors in GI images has no quantitative analysis in detail because of limited pages. Therefore, in this paper, we completely propose a series of mathematical statistics to systematically analyze the color sensitivity in GI images from the RGB color space domain to the 2D DCT spatial frequency domain in order to make up for a deficiency in our previous work [21]. This paper also refines the experimental results to analyze the performance about the compression rate, the quality degradation, and the ability of power saving individually.

As per the analysis of color sensitivity, the sensitivity of GI image sharpness to red component is at the same level as the sensitivity to green component. This result shows that the GI image is cardinal and different from the general image, whose sharpness sensitivity to the green component is much higher than the sharpness sensitivity to the red component. Because the GICam-II starts compressing the GI image from the Bayer-patterned image, the GICam-II technique subsamples the green component to make the weighting of red and green components the same. Besides, since the sharpness sensitivity to the blue component is as low as 7%, the blue component is downsampled by four. As shown in experimental results, with the compression ratio as high as 4 : 1, the GICam-II can significantly save the power dissipation by 38.5% when compared with previous GICam work [14] and 98.95% when compared with JPEG compression, while the average PSNRY is 40.73 dB. The rest of the paper is organized as follows. Section 2 introduces fundamentals of GICam compression and briefs the previous GICam work. Section 3 presents the sensitivity analysis of GICam image and shows the importance of red component in GI image. In Section 4, the GICam-II compression will be described in detail. Then, Section 5 illustrates the experimental results in terms of compression ratio, image quality, and power consumption. Finally, Section 6 concludes our contribution and merits of this work.

2. The Review of GICam Image Compression Algorithm

Instead of applying state-of-the-art video compression techniques, we proposed a simplified image compression algorithm, called GICam. Traditional compression algorithms employ the YCbCr quantization to earn a good compression ratio while the visual distortion is minimized, based on the factors related to the sensitivity of the human visual system (HVS). However, for the sake of power saving, our compression rather uses the RGB quantization [22] to save the computation of demosaicing and color space transformation. As mentioned above, the advantage of applying RGB quantization is twofold: saving the power dissipation on preprocessing steps and reducing the computing load of 2D DCT and quantization. Moreover, to reduce the hardware cost and quantization power dissipation, we have modified the RGB quantization tables, and the quantization multipliers are the power of two. In GICam, the Lempel-Ziv (LZ) coding [23] is employed for the entropy coding. The reason we adopted LZ coding as the entropy coding is because the LZ encoding does not need look-up tables and complex computation. Thus, the LZ encoding consumes less power and uses smaller silicon size than the other candidates, such as the Huffman encoding and the arithmetic coding. The target compression performance of the GICam image compression is to reduce image size by at least 75%. To meet the specification, given the quantization tables, we exploited the cost-optimal LZ coding parameters to meet the compression ratio requirement by simulating with twelve tested endoscopic pictures shown in Figure 4.

When comparing the proposed image compression with the traditional one in [14], the power consumption of GICam image compressor can save 98.2% because of the reduction of memory requirement. However, extending the utilization of battery life for a capsule endoscope remains an important issue. The memory access dissipates the most power in GICam image compression. Therefore, in order to achieve the target of extending the battery life, it is necessary to consider how to efficiently reduce the memory access.

3. Analysis of Sharpness Sensitivity in Gastrointestinal Images

3.1. The Distributions of Primary Colors in the RGB Color Space. In the modern color theory [24, 25], most color spaces in use today are oriented either toward hardware design or toward product applications. Among these color spaces, the RGB (red, green, blue) space is the most commonly used in the category of digital image processing,

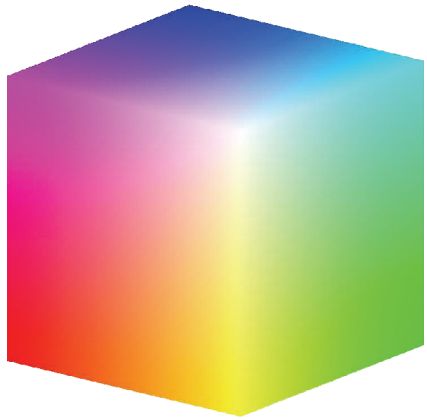


Figure 1: The RGB color space.

especially, broad class of color video cameras, and we consequently adopt the RGB color space to analyze the importance of primary colors in the GI images. In the RGB color space, each color appears in its primary spectral components of red, green, and blue. The RGB color space is based on a Cartesian coordinate system and is the cube shown in Figure 1 in which, the differ colors of pixels are points on or inside the cube based on the triplet of values (R, G, B). The block-based image data can be sequentially outputted via the proposed locally raster-scanning mechanism for this raw image sensor. The reason for adopting a novel image sensor without using generally conventional ones is to efficiently save the size of buffer memory. Conventional raw image sensors adopt the raster-scanning mechanism to output the image pixels sequentially, but they need large buffer memory to form each block-based image data before executing the block-based compression. However, we only need a small ping-pong type memory structure to directly save the block-based image data from the proposed locally raster-scanning raw image sensor. The structure of this raw image sensor is shown in Figure 2(a), and the pixel sensor architecture for the proposed image sensor is shown in Figure 2(b). In order to prove the validity for this novel image sensor before the fabrication via the Chung-Shan Institute of Science and Technology, the chip of the 32-by-32 locally raster-scanning raw image sensor was designed by full-custom CMOS technology, and this chip is submitted to Chip Implementation Center (CIC), Taiwan, for the fabrication. Figures 3(a) and 3(b), respectively, show the chip layout and the package layout with the chip specification. The advantage of this novel CMOS image sensor can save the large area of buffer memory. The size of buffer memory can be as a simple ping-pong memory structure shown in Figure 10 while executing the proposed image algorithm, a novel block coding.

Our research only focuses on developing the proposed image compressor, and other components are implemented by another research department for the GICam-II capsule endocopy. Therefore, the format of the GI image used in the simulation belongs to a raw image from the 512-by-512 sensor designed by Chung-Shan Institute of Science and Technology. In this work, we applied twelve GI images captured shown in Figure 4 for test cases to evaluate the compression technique. The distribution of GI image pixels in the RGB color space is nonuniform. Obviously, the GI image is reddish, and the pixels are amassed to the red region. Based on the observation in the RGB color space, the majority of red values are distributed between 0.5 and 1 while most of the green and blue values are distributed between 0 and 0.5 for all tested GI images.

To further analyze the chrominance distributions and variations in the RGB color space for each tested GI image, two quantitative indexes are used to quantify these effects. The first index is to calculate the average distances between total pixels and the maximum primary colors in each GI image, and the calculations are formulated as (1), (2), and (3). First, (1) defines the average distance between total pixels and the most red color $(\overline{R})$, in which $R(i, j)$ means the value of red component of one GI image at (i, j) position and the value of most red color $(R_{\max})$ is 255. In addition, M and N represent the width and length for one GI image, respectively. The M is 512, and the N is 512 for twelve tesed GI images in this work. Next, (2) also defines the average distance between total pixels and the most green color $(\overline{G})$, and the value of the most green one $(G_{\max})$ is 255. Finally, (3) defines the average distance between total pixels and the most blue color $(\overline{B})$, and the value of the most blue color $(B_{\max})$ is 255. Table 1 shows the statistical results of $\overline{R}$, $\overline{G}$, and $\overline{B}$ for all tested GI images. From Table 1, the results clearly show that $\overline{R}$ has the shortest average distance. Therefore, human eyes can be very sensitive to the obvious cardinal ingredient on all surfaces of tested GI images. Moreover, comparing $\overline{G}$ with $\overline{B}$, $\overline{G}$ is shorter than $\overline{B}$ because $\overline{G}$ contributes larger proportion in luminance.

We have

$$\overline{R} = E\left[\left(1 - \frac{R(i,j)}{R_{\max}}\right)\right] = \left(\frac{1}{M \times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1 - \frac{R(i,j)}{R_{\max}}\right), \tag{1}$$

$$\overline{G} = E\left[\left(1 - \frac{G(i,j)}{G_{\max}}\right)\right] = \left(\frac{1}{M \times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1 - \frac{G(i,j)}{G_{\max}}\right), \tag{2}$$

$$\overline{B} = E\left[\left(1 - \frac{B(i,j)}{B_{\max}}\right)\right] = \left(\frac{1}{M \times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1 - \frac{B(i,j)}{B_{\max}}\right). \tag{3}$$

The first index has particularly quantified the chrominance distributions through the concept of average distance, and the statistical results have also shown the reason the human eyes can sense the obvious cardinal ingredient for all tested GI images. Next, the second index is to calculate the variance between total pixels and average distance, in order to further observe the color variations in GI images, and

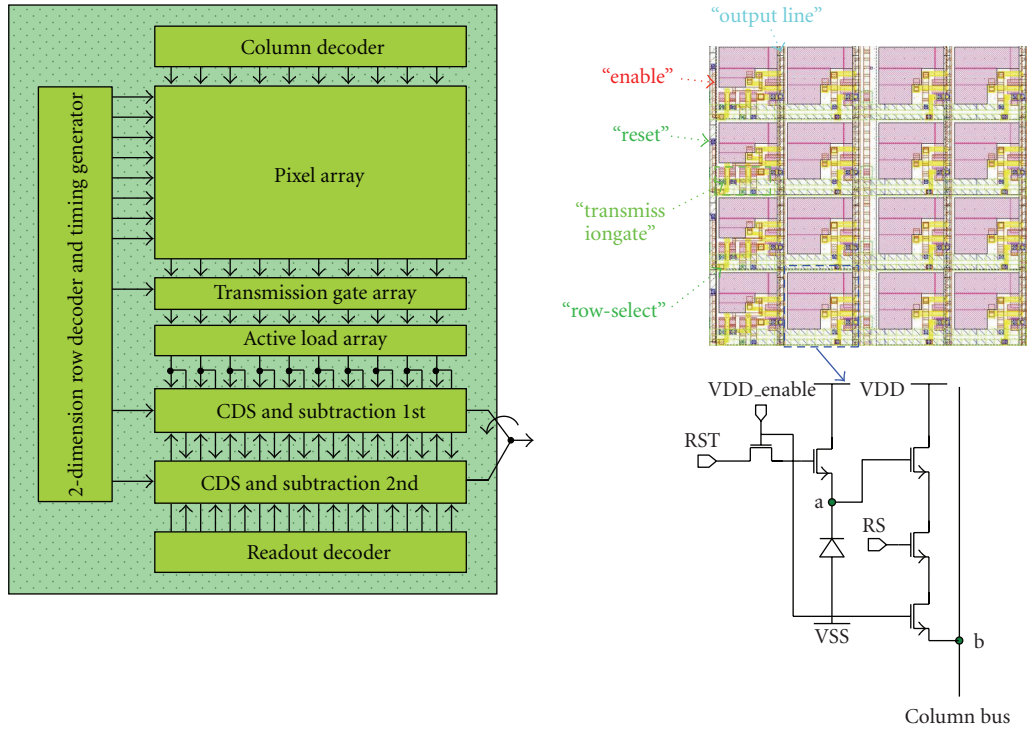


FIGURE 2: (a) The structure of locally raster-scanning raw image sensor. (b) The pixel sensor architecture for the locally raster-scanning raw image sensor.

the calculations are formulated as (4). Table 2 shows that the average variation of red signal is 0.09, the average variance of green one is 0.03, and the average variance of blue one is 0.02. It signifies that the color information of red signal must be preserved carefully more than the other two primary colors, green and blue, for GI images because the dynamic range of red signal is broader than the green and blue ones. In addition, the secondary is green signal, and the last is blue signal.

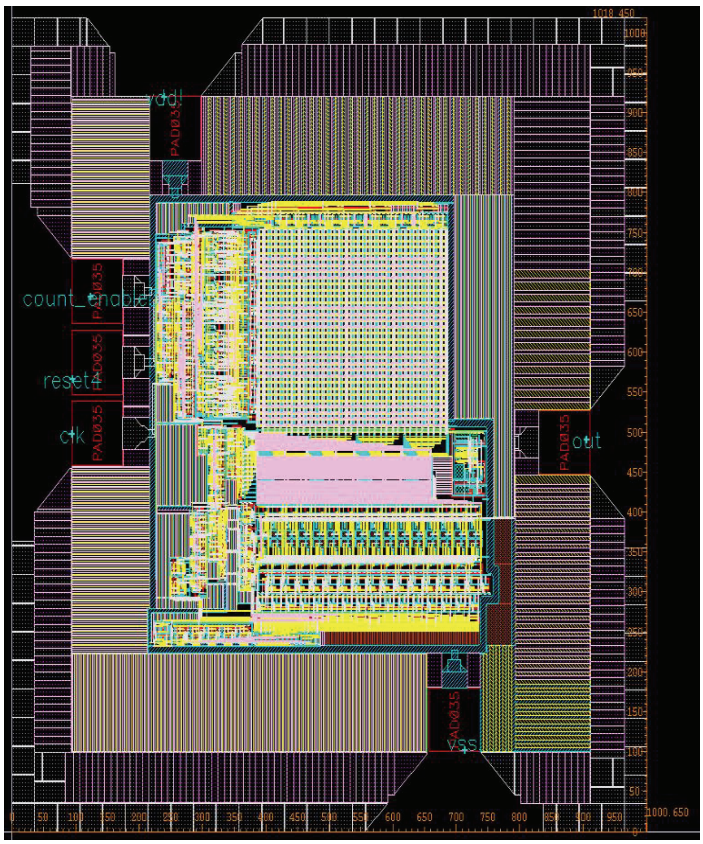
We have

$$\begin{aligned}
\mathrm{VAR}_R &= E\left[\left(1-\frac{R(i,j)}{R_{\max}}\right)^2\right]-\left\{E\left[\left(1-\frac{R(i,j)}{R_{\max}}\right)\right]\right\}^2\\
&=\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[1-\frac{R(i,j)}{R_{\max}}\right]^2\\
&\quad-\left[\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1-\frac{R(i,j)}{R_{\max}}\right)\right]^2,\\
\mathrm{VAR}_G &= E\left[\left(1-\frac{G(i,j)}{G_{\max}}\right)^2\right]-\left\{E\left[\left(1-\frac{G(i,j)}{G_{\max}}\right)\right]\right\}^2\\
&=\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[1-\frac{G(i,j)}{G_{\max}}\right]^2\\
&\quad-\left[\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1-\frac{G(i,j)}{G_{\max}}\right)\right]^2,\\
\mathrm{VAR}_B &= E\left[\left(1-\frac{B(i,j)}{B_{\max}}\right)^2\right]-\left\{E\left[\left(1-\frac{B(i,j)}{B_{\max}}\right)\right]\right\}^2\\
&=\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[1-\frac{B(i,j)}{B_{\max}}\right]^2\\
&\quad-\left[\left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1-\frac{B(i,j)}{B_{\max}}\right)\right]^2.
\end{aligned}\tag{4}$$

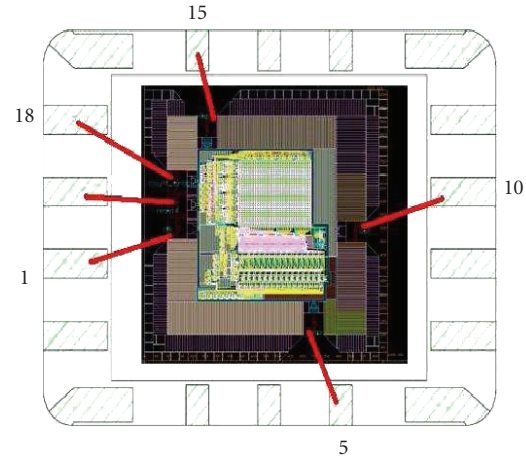
3.2. The Analysis of Sharpness Sensitivity to Primary Colors for Gastrointestinal Images. Based on the analysis of RGB color space, the importance of chrominance is quantitatively demonstrated for GI images. Except for the chrominance, the luminance is another important index because it can efficiently represent the sharpness of an object. Equation (5) is the formula of luminance (Y) and the parameters $a1$, $a2$, and $a3$ are 0.299, 0.587, and 0.114, respectively:

$$Y = a1\times R + a2\times G + a3\times B. \tag{5}$$

To efficiently analyze the importance of primary colors in the luminance, the analysis of sensitivity is applied. Through the analysis of sensitivity, the variation of luminance can actually reflect the influence of each primary colors. Equation (6) defines the sensitivity of red ($S_{Y_{i,j}}^{R_{i,j}}$), the sensitivity of green



(a)



Technology	0.35 μm
Voltage	3.3 V
Sensor array size	32-by-32
Power consumption	8.8586 mW
Chip size	1.000651.01845 mm^2
Output	Analog output

(b)

Figure 3: (a) The chip layout of the locally raster-scanning raw image sensor. (b) The package layout and the chip specification of the locally raster-scanning raw image sensor.

Table 1: The analysis of average distance.

	Average distance		
Test picture ID	$\overline{R}$	$\overline{G}$	$\overline{B}$
1	0.58	0.80	0.82
2	0.55	0.74	0.79
3	0.54	0.81	0.86
4	0.55	0.76	0.81
5	0.66	0.82	0.85
6	0.66	0.84	0.87
7	0.59	0.82	0.88
8	0.68	0.81	0.83
9	0.55	0.80	0.85
10	0.53	0.81	0.84
11	0.53	0.81	0.86
12	0.62	0.80	0.85
Average	0.59	0.80	0.84

Table 2: The analysis of variance.

	Variance of distance		
Test picture ID	VAR_R	VAR_G	VAR_B
1	0.08	0.02	0.02
2	0.11	0.05	0.03
3	0.10	0.03	0.02
4	0.10	0.04	0.02
5	0.07	0.02	0.01
6	0.08	0.02	0.01
7	0.09	0.02	0.01
8	0.06	0.02	0.02
9	0.09	0.03	0.01
10	0.10	0.03	0.02
11	0.10	0.03	0.02
12	0.10	0.04	0.02
Average	0.09	0.03	0.02

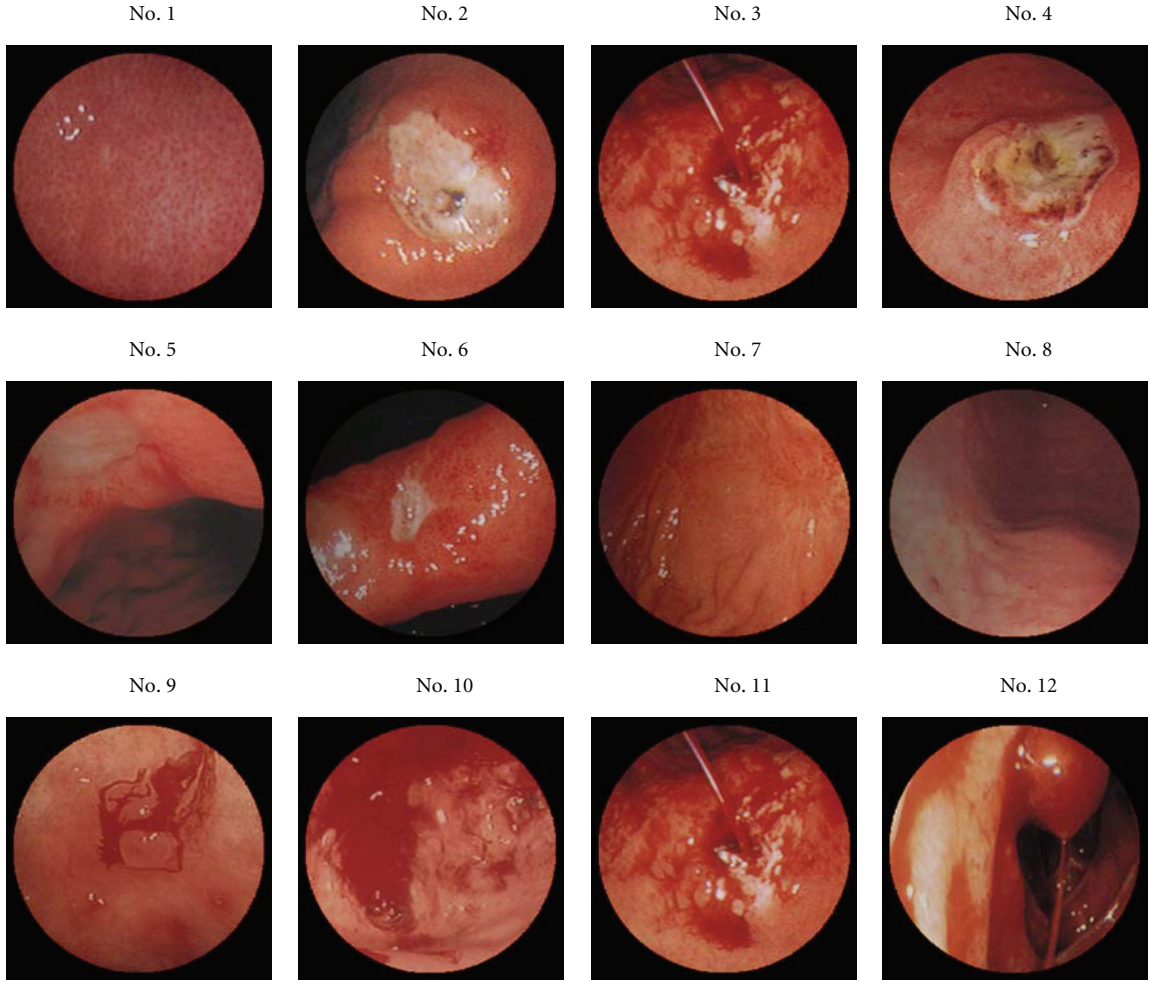


Figure 4: The twelve tested GI images.

($S_{Y_{i,j}}^{G_{i,j}}$), and the sensitivity, of blue ($S_{Y_{i,j}}^{B_{i,j}}$) at position (i, j), respectively for a color pixel of a GI image:

$$\begin{aligned} S_{Y_{i,j}}^{R_{i,j}} &= \frac{\Delta Y_{i,j}/Y_{i,j}}{\Delta R_{i,j}/R_{i,j}} = \frac{R_{i,j}}{Y_{i,j}} \times \frac{\Delta Y_{i,j}}{\Delta R_{i,j}} = \frac{a1 \times R_{i,j}}{Y_{i,j}}, \\ S_{Y_{i,j}}^{G_{i,j}} &= \frac{\Delta Y_{i,j}/Y_{i,j}}{\Delta G_{i,j}/G_{i,j}} = \frac{G_{i,j}}{Y_{i,j}} \times \frac{\Delta Y_i}{\Delta G_{i,j}} = \frac{a2 \times G_{i,j}}{Y_{i,j}}, \\ S_{Y_{i,j}}^{B_{i,j}} &= \frac{\Delta Y_{i,j}/Y_{i,j}}{\Delta B_{i,j}/B_{i,j}} = \frac{B_{i,j}}{Y_{i,j}} \times \frac{\Delta Y_i}{\Delta B_{i,j}} = \frac{a3 \times B_{i,j}}{Y_{i,j}}. \end{aligned} \tag{6}$$

After calculating the sensitivity of each primary color for a GI image, the average sensitivity of red ($\overline{S_Y^R}$), the average sensitivity of green ($\overline{S_Y^G}$), and the average sensitivity of blue ($\overline{S_Y^B}$) are calculated by (7) for each GI image. M and N represent the width and length for a GI image, respectively. Table 3 shows the average sensitivities of red, green, and blue for all tested GI images. From the calculational results, the sensitivity of blue is the slightest, and hence the variation of luminance arising from the aliasing of blue is very invisible. In addition to the sensitivity of blue, the sensitivity of red is close to the one of green, and thus they both have a very close influence on the variation of luminance.

We have

$$\begin{aligned} \overline{S_Y^R} &= \left(\frac{1}{M \times N}\right) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} S_{Y_{i,j}}^{R_{i,j}}, \\ \overline{S_Y^G} &= \left(\frac{1}{M \times N}\right) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} S_{Y_{i,j}}^{G_{i,j}}, \\ \overline{S_Y^B} &= \left(\frac{1}{M \times N}\right) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} S_{Y_{i,j}}^{B_{i,j}}. \end{aligned} \tag{7}$$

To sum up the variance of chrominance and the sensitivity of luminance, blue is the most insensitive color in the GI images. Therefore, the blue component can be further downsampled without significant sharpness degradation. Moreover, comparing the red signal with the green signal, they both have a very close influence on the variation of luminance, because they have very close sensitivities. However, the chrominance of red varies more than the chrominance of green, and hence the information completeness of red has higher priority than the green. Because the proposed compression coding belongs to the DCT-based image coding, the coding is processed in the spatial-frequency domain. To let the priority relationship between red and green also response in the spatial-frequency domain,

Table 3: The analysis of average sensitivities.

	The sensitivity of primary colors in luminance		
Test picture ID	$\overline{S_Y^R}$	$\overline{S_Y^G}$	$\overline{S_Y^B}$
1	0.49	0.43	0.08
2	0.44	0.48	0.08
3	0.55	0.39	0.06
4	0.47	0.46	0.07
5	0.45	0.47	0.08
6	0.48	0.45	0.07
7	0.52	0.42	0.06
8	0.44	0.48	0.08
9	0.51	0.43	0.06
10	0.54	0.40	0.06
11	0.55	0.39	0.06
12	0.49	0.44	0.07
Average	0.49	0.44	0.07

the analysis of alternating current (AC) variance will be accomplished to demonstrate the inference mentioned above in the next subsection.

3.3. The Analysis of AC Variance in the 2D DCT Spatial Frequency Domain for Gastrointestinal Images. According to the analysis results from the distributions of primary colors in the RGB color space and the proportion of primary colors in the luminance for GI images, the red signal plays a decisive role in the raw image. The green signal plays a secondary role, and the blue signal is very indecisive. To verify the validity of observation mentioned above, we first use the two-dimensional (2D) 8×8 discrete cosine transform (DCT) to transfer the spatial domain into the spatial-frequency domain for each of the components, R, G1, G2, and B. The 2D 8×8 DCT transformation can be perceived as the process of finding for each waveform in the 2D 8×8 DCT basic functions and also can be formulated as (8) for each 8×8 block in R, G1, G2, and B subimages, respectively. M and N represent the width and length for one GI image, respectively. k, $l = 0, 1, \ldots, 7$, and y_{kl} is the corresponding weight of DCT basic function in the kth row and the lth column. P represents the total number of pictures and B represents the total number of 8×8 blocks in the GI images.

We have

$$R_{pb}(kl) = \frac{c(k)}{2}\sum_{i=0}^{7}\left[\frac{c(l)}{2}\sum_{j=0}^{7} r_{ij}\cos\left(\frac{(2j+1)l\pi}{16}\right)\right] \times \cos\left(\frac{(2i+1)k\pi}{16}\right),$$

$$G_{pb}(kl) = \frac{c(k)}{2}\sum_{i=0}^{7}\left[\frac{c(l)}{2}\sum_{j=0}^{7} g_{ij}\cos\left(\frac{(2j+1)l\pi}{16}\right)\right] \times \cos\left(\frac{(2i+1)k\pi}{16}\right),$$

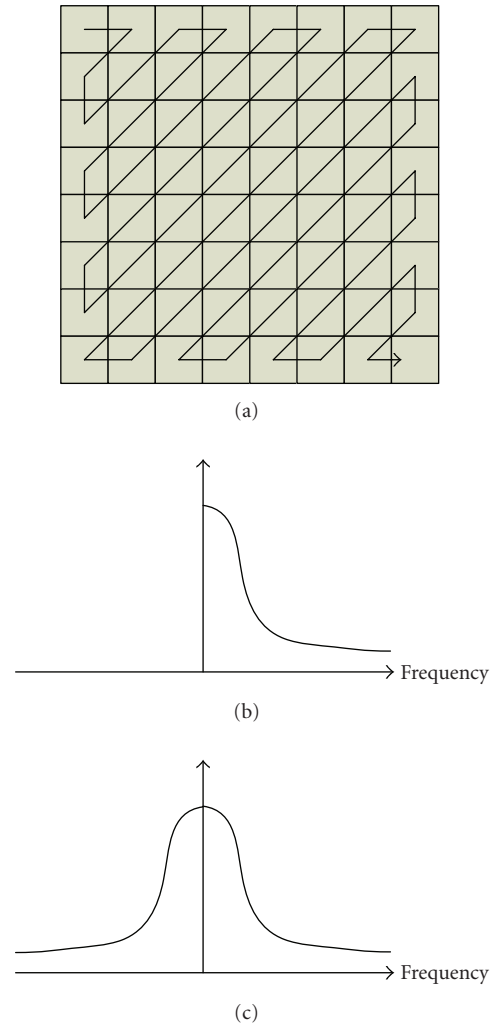


Figure 5: (a) Zigzag scanning for a 8×8 block (b) 1D signal distribution after zigzag scanning order. (c) The symmetric type of frequency for the 1D signal distribution.

$$B_{pb}(kl) = \frac{c(k)}{2}\sum_{i=0}^{7}\left[\frac{c(l)}{2}\sum_{j=0}^{7} b_{ij}\cos\left(\frac{(2j+1)l\pi}{16}\right)\right] \times \cos\left(\frac{(2i+1)k\pi}{16}\right),$$

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } k = 0, \\ 1, & \text{otherwise,} \end{cases}$$

$$c(l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } l = 0, \\ 1 & \text{otherwise,} \end{cases} \quad (8)$$

Next, we calculate the average energy amplitude of all alternating current (AC) coefficients of all tested GI images, in order to observe the variation of energy for each of the components R, G1, G2, and B, and the calculations are

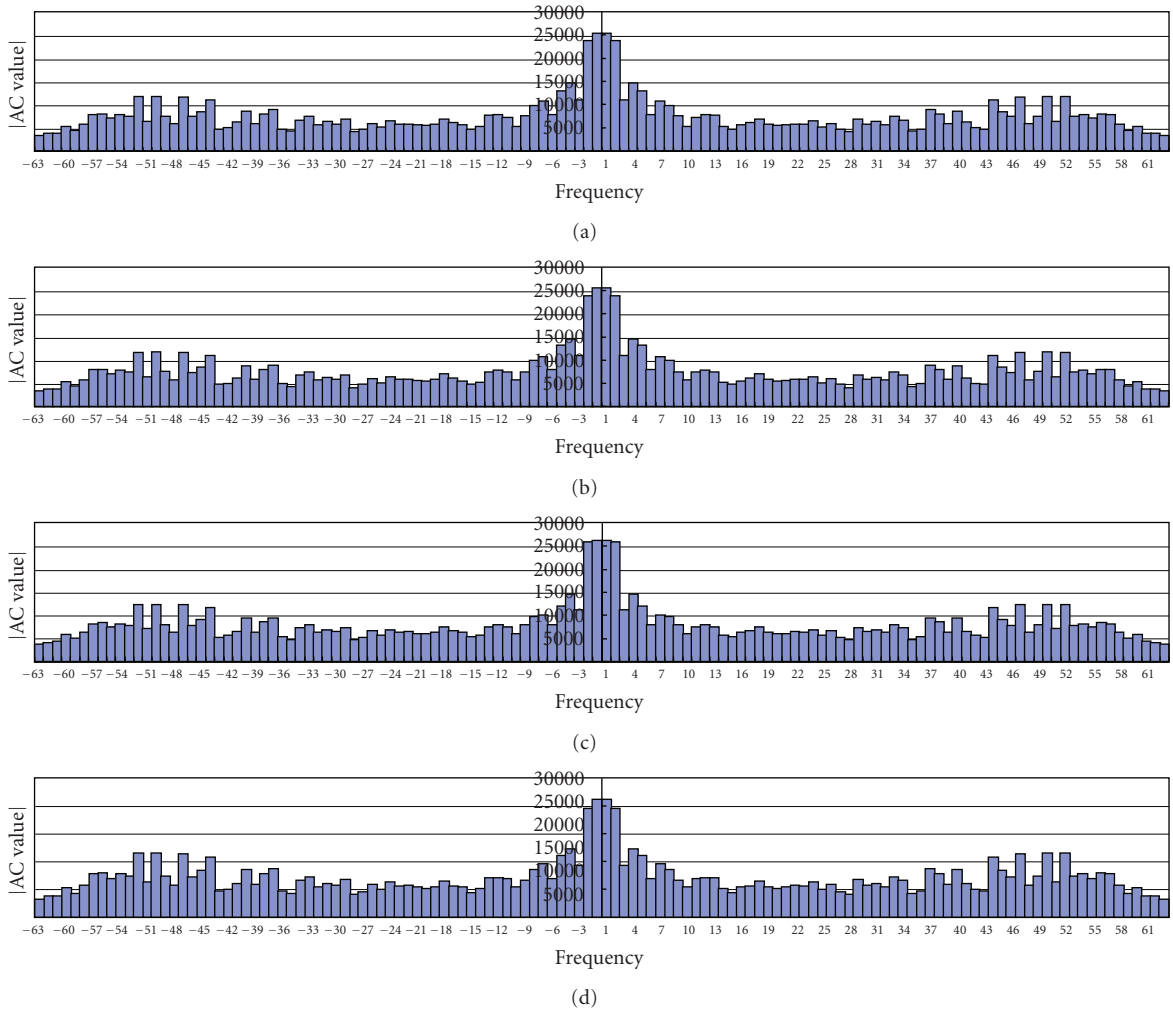


Figure 6: (a) Spatial-frequency distribution converting into one dimension for G1 component. (b) Spatial-frequency distribution converting into one dimension for G2 component. (c) Spatial-frequency distribution converting into one dimension for R component. (d) Spatial-frequency distribution converting into one dimension for B component.

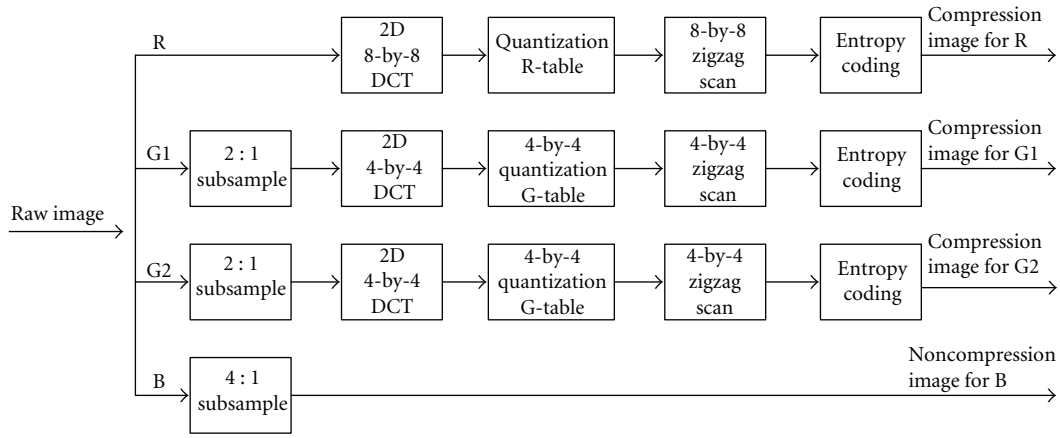


Figure 7: The GICam-II image compression algorithm.

formulated as

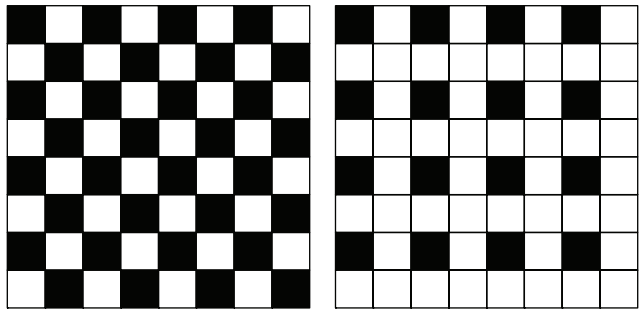
$$A_R(kl) = \frac{1}{P}\sum_{p=1}^{P}\left[\sum_{b=0}^{B-1}\left|R_{pb}(kl)\right|\right],$$
$$A_G(kl) = \frac{1}{P}\sum_{p=1}^{P}\left[\sum_{b=0}^{B-1}\left|G_{pb}(kl)\right|\right], \quad (9)$$
$$A_B(kl) = \frac{1}{P}\sum_{p=1}^{P}\left[\sum_{b=0}^{B-1}\left|B_{pb}(kl)\right|\right].$$

After calculating the average energy amplitude, we convert the 2D DCT domain into one-dimensional (1D) signal distribution in order to conveniently observe the variation of frequency. Consequently, a tool for transforming two-dimensional signals into one dimension is needed. There are many schemes to convert 2D into 1D, including row-major scan, column-major scan, peano-scan, and zigzag scan. Majority of the DCT coding schemes adopt zigzag scan to accomplish the goal of conversion, and we use it here. The benefit of zigzag is its property of compacting energy to low-frequency regions after discrete cosine transformation. The arrangement sorts the coefficients from low to high frequency, and Figure 5(a) shows the zigzag scanning order for 8×8 block. Figure 5(b) shows the 1D signal distribution after zigzag scanning order, and Figure 5(c) shows the symmetric type of frequency for the 1D signal distribution.

Through the converting method of Figure 5, the 1D signal distributions of each R, G1, G2, B component are shown in Figure 6. The variances of frequency are 1193, 1192, 1209, and 1244 for G1, G2, R, and B, respectively, and the variance of R is very close to the ones of G1 and G2 from the result. However, the data of G are twice the data of R based on the Bayer pattern and hence, the data of G can be reduced to half at the most. Based on the analysis result mentioned above, the R component is very decisive for GI images, and it needs to be compressed completely. However, the G1, G2, and B components do not need to be compressed completely because they are of less than the R component. Therefore, in order to efficiently reduce the memory access to expend the battery life of capsule endoscopy, the data of G1, G2, and B components should be appropriately decreased according to the proportion of their importance prior to the compression process. In this paper, we successfully propose a subsample-based GICam image compression algorithm, and the proposed algorithm firstly uses the subsample technique to reduce the incoming data of G1, G2, and B components before the compression process. The next section will describe the proposed algorithm in detail.

4. The Subsample-Based GICam Image Compression Algorithm

Figure 7 illustrates the GICam-II compression algorithm. For a 512×512 raw image, the raw image firstly divides into four parts, namely, R, G1, G2, and B components and each of the components has 256×256 pixels. For the R component, the incoming image size to the 2D DCT is $256 \times 256 \times 8$ bits, where the incoming image data are completely compressed because of the importance itself in GI images.



(a) (b)

Figure 8: (a) 2 : 1 subsample pattern. (b) 4 : 1 subsample pattern.

Except for the R component, the GICam-II algorithm can use an appropriate subsample ratio to pick out the necessary image pixels into the compression process for G1, G2, and B components, and (10) and (11) are formulas for the subsample technique. $\mathrm{SM}_{16:2m}$ is the subsample mask for the subsample ratio 16-to-2m as shown in (10), and the subsample mask $\mathrm{SM}_{16:2m}$ is generated from basic mask as shown in (11). The type of subsample direction is block-based, when certain positions in the subsample mask are one, their pixels in the same position will be compressed, or otherwise they are not processed. For the G1 and G2 components, the low subsample ratio must be assigned, considering their secondary importance in GI images. Thus, the 2 : 1 subsample ratio is candidate one, and the subsample pattern is shown in Figure 8(a). Finally, for the B component, the 4 : 1 subsample ratio is assigned, and the subsample pattern is shown in Figure 8(b). In the GICam-II image compression algorithm, the 8×8 2D DCT is still used to transfer the R component. However, the 4×4 2D DCT is used for G1 and G2 components because the incoming data are reduced by subsample technique. Moreover, the G quantization table is also modified and shown in Figure 9. Finally, the B component is directly transmitted, not compressed, after extremely decreasing the incoming data. Because of the noncompression for the B component, the 8×8 and 4×4 zigzag scanning techniques are added into the GICam-II to further increase the compression rate for R, G1, and G2 components before entering the entropy encoding. In the GICam-II, the Lempel-Ziv (LZ) coding [23] is also employed for the entropy coding because of nonlook-up tables and low complex computation.

We have

$$\mathrm{SM}_{16:2m}(i,j) = \mathrm{BM}_{16:2m}(i \bmod 4, j \bmod 4) \quad m = 1,2,3,4,5,6,7,8, \quad (10)$$

$$\mathrm{BM}_{16:2m}(k,l) = \begin{bmatrix} u(m-1) & u(m-5) & u(m-2) & u(m-6) \\ u(m-7) & u(m-3) & u(m-8) & u(m-4) \\ u(m-2) & u(m-5) & u(m-1) & u(m-6) \\ u(m-7) & u(m-3) & u(m-8) & u(m-4) \end{bmatrix},$$

where $u(n)$ is a step function,

$$u(n) = \begin{cases} 1, & \text{for } n \geq 0 \\ 0, & \text{for } n < 0. \end{cases} \quad (11)$$

5. The Architecture of Subsample-Based GICam Image Compressor

Figure 10 shows the architecture of the GICam-II image compressor, and it faithfully executes the proposed GICam-II image compression algorithm shown in Figure 7. The window size, w, and the maximum matching length, l, parameters for LZ77 encoder can be loaded into the parameter register file via a serial interface after the initial setting of the hardware reset. Similarly, coefficients of 2D DCT and parameters of initial setting for all controllers shown in Figure 10 can be also loaded into the parameter register file. The GICam-II image compressor processes the image in the block order of G1, R, G2, and B. Because the data stream from the image sensor is block based, the GICam-II image compressor adopts the structure of ping-pong memory to hold each block of data. The advantage of using this structure is the high parallelism between the data loading and data processing.

When the GICam-II image compressor begins, the proposed architecture first loads the incoming image in the block order of G1, R, G2, and B from the image sensor and passes them with the valid signal control via the Raw-Data Sensor Interface. The Raw-Data Sensor Interface is a simple register structure with one clock cycle delay. This design absolutely makes sure that no any glue-logic circuits that can affects the timing of logic synthesis exists between the raw image sensor and the GICam-II image compressor. The Downsample Controller receives the valid data and then selects the candidate subsample ratio to sample the candidate image data in the block order of G1, R, G2, and B. The Ping-Pong Write Controller can accurately receive the data loading command from the Downsample Controller and then push the downsample image data into the candidate one of the ping-pong memory. At the same time, the Ping-Pong Read Controller pushes the stored image data from another memory into the Transformation Coding. The Ping-Pong Write Controller and the Ping-Pong Read Controller will issue an announcement to the Ping-Pong Switch Controller, respectively, while each data access is finished. When all announcement arrives in turn, the Ping-Pong Switch Controller will generate a pulse-type Ping-Pong Switching signal, one clock cycle, to release each announcement signal from the high level to zero for the Ping-Pong Write Controller and the Ping-Pong Read Controller. The Ping-Pong Switch Counter also uses the Ping-Pong Switching signal to switch the read/write polarity for each memory in the structure of the Ping-Pong Memory.

The Transformation Coding consists of the 2D DCT and the quantizer. The goal of the transformation coding is to transform processing data from the spatial domain into the spatial frequency domain and further to shorten the range in the spatial frequency domain before entropy coding in order to increasing the compression ratio. The 2D DCT alternatively calculates row or column 1D DCTs. The 1D DCT is a multiplierless implementation using the algebraic integer encoding [14]. The algebraic integer encoding can minimize the number of addition operations. As regards the RG quantizer, the GICam-II image compressor utilizes the barrel shifter for power-of-two products. The power-of-two quantization table shown in Figure 9 can reduce the cost of multiplication while quality degradation is quite little. In addition, the 8-by-8 memory array between the quantizer and the LZ77 encoder is used to synchronize the operations of quantization and LZ77 encoding. Since the frame rate of GICam-II image compressor is 2 frames/second, the 2D DCT can be folded to trade the hardware cost with the computing speed, and the other two data processing units, quantization and LZ77 encoder, can operate at low data rate. Due to noncompression for the B component, the B component is directly transmitted from the ping-pong memory, not compressed. Finally, the LZ77 encoder is implemented by block-matching approach and the details of each processing element and overall architecture have been also shown in [14].

6. Experimental Results

We have particularly introduced the method of efficiently decreasing the incoming data with the subsample technique in the GICam-II compression algorithm. The performance of the compression rate, the quality degradation, and the ability of power saving will then be experimentally analyzed using the GICamm-II compressor.

6.1. The Analysis of Compression Rate for GI Images. In this paper, twelve GI images are tested and shown in Figure 4. First of all, the target compression performance of the GICam-II image compression is to reduce image size by at least 75%. To meet the specification, we have to exploit the cost-optimal LZ coding parameters. There are two parameters in the LZ coding to be determined: the window size, w, and the maximum matching length, l. The larger the parameters are, the higher the compression ratio will be; however, the implementation cost will be higher. In addition, there are two kinds of LZ codings in the GICam-II compressor; one is $R(w,l)$ for R component, and the other is $G(w,l)$ for G1 and G2 components. We set the values of parameters by using a compression ratio of $4:1$ as the threshold. Our goal is to determine the minimum $R(w,l)$ and $G(w,l)$ sets under the constraint of $4:1$ compression ratio.

The compression ratio (CR) is defined as the ratio of the raw image size to the compressed image size and formulated as (12). The measure of the compression ratio is the compression rate. The formula of the compression rate is calculated by (13). The results in Figure 11 are shown by simulating the behavior model of GICam-II compressor; it is generated by MATLAB. As seen in Figure 11, simulating with twelve endoscopic pictures, (32, 32) and (16, 8) are the minimum $R(w,l)$ and $G(w,l)$ sets to meet the compression ratio requirement. The subsample technique of the GICam-II compressor initially reduces the input image size by $43.75\%\ ((1-1/4-(1/4*1/2*2)-(1/4*1/4))*100\%)$ before executing the entropy coding, LZ77 coding. Therefore, the overall compression ratio of GICam-II compressor minus 43.75% is the compression effect of LZ77 coding that

32	32	32	32	32	32	64	64
32	16	16	32	32	64	64	128
32	16	16	32	32	64	128	128
32	32	32	32	64	64	128	256
32	32	32	64	64	128	128	256
64	64	64	128	128	128	256	256
64	128	128	128	256	256	256	256
128	128	128	256	256	256	256	512

(a)

16	16	32	32
16	16	32	64
32	32	64	64
64	64	128	128

(b)

Figure 9: (a) The modified R quantization table. (b) The modified G quantization table.

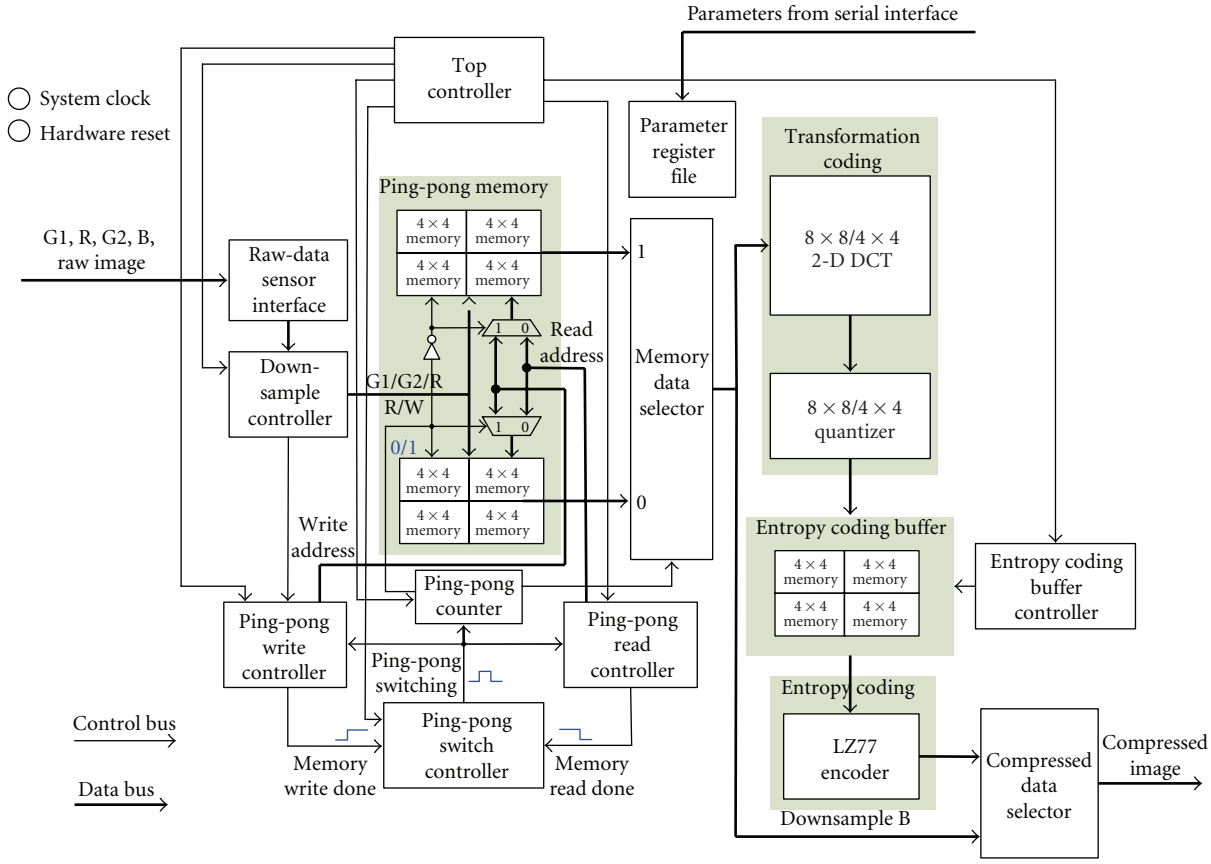


Figure 10: The GICam-II image compressor.

combines with the quantization, and the simulation results are shown in Figure 12.

This research paper focuses on proposing a subsample-based low-power image compressor for capsule gastrointestinal endoscopy. This obvious reddish characteristic is due to the slightly luminous intensity of LEDs and the formation of image in the capsule gastrointestinal endoscopy. The GICam-II compression algorithm is motivated on the basis of this reddish pattern. Therefore, we do not consider compressing other endoscopic images except for gastrointestinal images to avoid the confusion of topic for this research. However, general endoscopic images generated via a wired endoscopic take on the yellow characteristic due to the vividly luminous intensity of LEDs. The yellow pattern mainly consists of red and green, and it also complies with the color sensitivity result in this research work. Therefore, I believe that the proposed GICam-II still supports good compression ratio for general endoscopic images.

We have

$$\text{Compression Ratio (CR)} = \frac{\text{bits before compression}}{\text{after compression}}, \tag{12}$$

$$\text{Compression Rate} = \left(1 - \text{CR}^{-1}\right) \times 100\%. \tag{13}$$

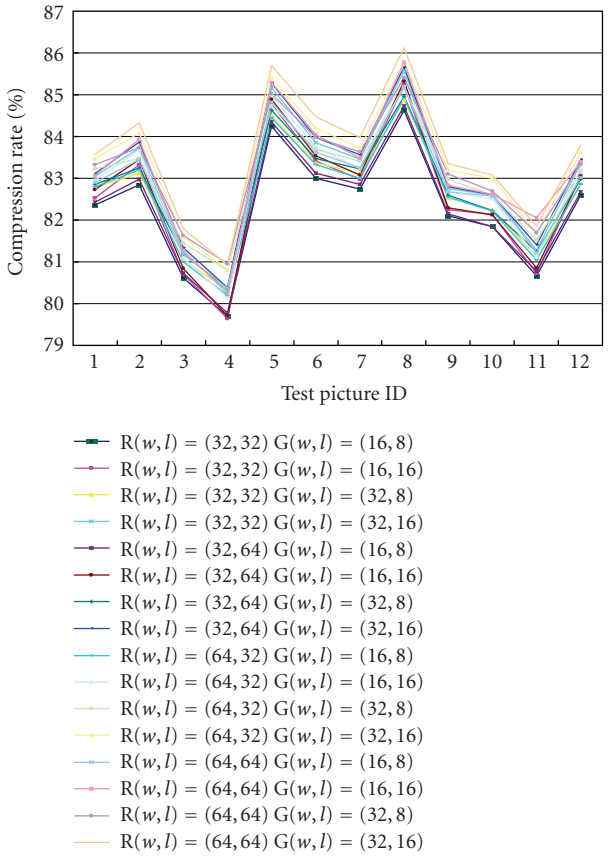


Figure 11: The compression performance of the GICam-II image compressor.

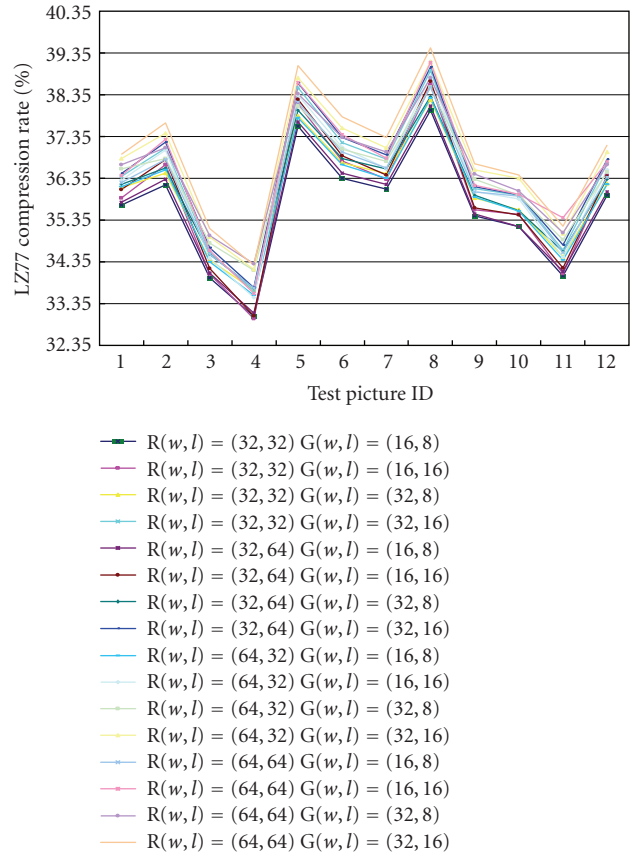


Figure 12: The compression performance of LZ77 coding that combines with the quantization in the GICam-II image compressor.

6.2. The Analysis of Compression Quality for GI Images. Using (32, 32) and (16, 8) as the parameter sets, in Table 4, we can see the performance in terms of the quality degradation and compression ratio. The measure of compression quality is the peak signal-to-noise ratio of luminance (PSNRY). The calculation of PSNRY is formulated as (14), where MSE is the mean square error of decompressed image and is formulated as (15). In (15), α_{ij} is the luminance value of original GI image, and β_{ij} is the luminance value of decompressed GI image. The result shows that the degradation of decompressed images is quite low while the average PSNRY is 40.73 dB. Figure 13 illustrates the compression quality of decoded test pictures. The difference between the original image and the decompressed image is invisible.

We have

$$\text{PSNRY} = 10\log_{10}\left(\frac{255^2}{\text{MSE}}\right), \tag{14}$$

$$\text{MSE} = \left(\frac{1}{M\times N}\right)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(\alpha_{ij}-\beta_{ij}\right)^2. \tag{15}$$

To demonstrate the validity of decompressed images, five professional gastroenterology doctors from the Division of Gastroenterology, Taipei Medical University Hospital, are invited to verify whether or not the decoded image quality is suitable for practical diagnosis. The criterion of evaluation is shown in Table 5. The score between 0 and 2 means that the diagnosis is affected, the score between 3 and 5 means that the diagnosis is slightly affected, and the score between 6 and 9 means that the diagnosis is not affected. According to the evaluation results of Figure 14, all decoded GI images are suitable for practical diagnosis because of high evaluation score, and the diagnoses are absolutely not affected, except for the 5th and 8th decoded images. The degrees of diagnoses are between no affection and extremely slight affection for the 5th and the 8th decoded images because only two doctors subjectively feel that their diagnoses are slightly affected. However, these two decoded images are not mistaken in diagnosis for these professional gastroenterology doctors. Therefore, the PSNRY being higher than 38 dB is acceptable according to the objective criterion of gastroenterology doctors.

6.3. The Analysis of Power Saving. To validate the GICam-II image processor, we used the FPGA board of Altera APEX 2100 K to verify the function of the GICam-II image processor, and the prototype is shown in Figure 15. After FPGA verification, we used the TSMC 0.18 μm 1P6M process to implement the GICam-II image compressor. When

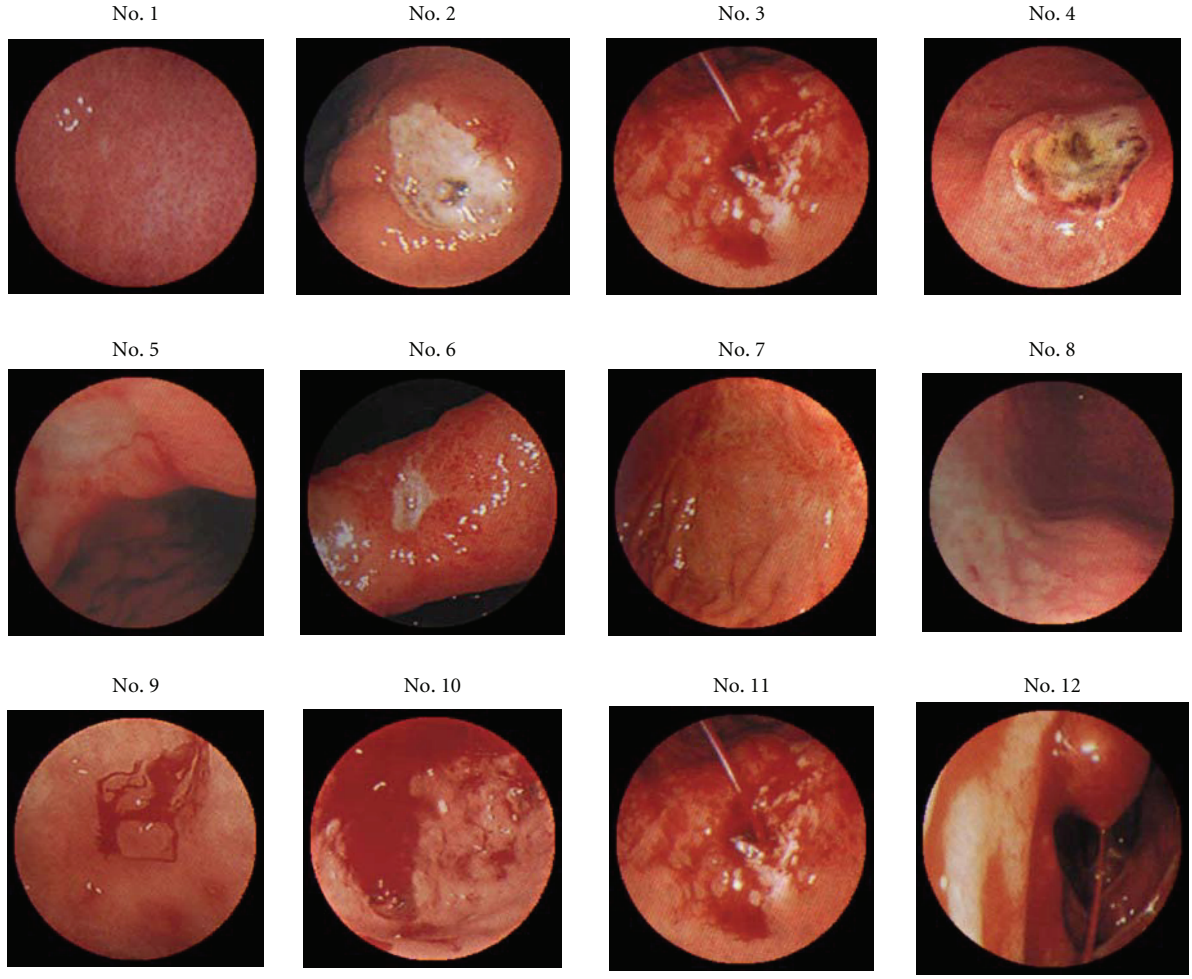


Figure 13: Demosaicked GI images.

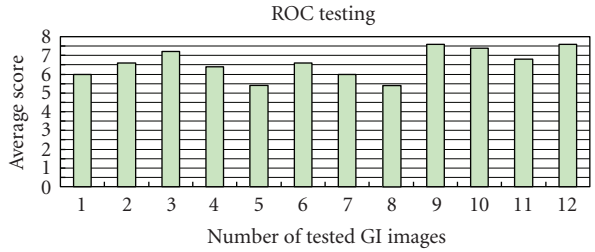


Figure 14: The evaluation results of professional gastroenterology doctors.

operating at 1.8 V, the power consumption of logic part is 3.88 mW, estimated by using PrimePower. The memory blocks are generated by Artisan memory compiler and consume 5.29 mW. The total power consumption is 9.17 mW for the proposed design. When comparing the proposed GICam-II image compressor with our previous GICam one in Table 6, the power dissipation can further save 38.5% under the approximate condition of quality degradation and compression ratio because of the reduction of memory requirement for G1, G2, and B components.

Table 4: The simulation results of twelve tested pictures.

Test picture ID	PSNRY (dB)	Compression rate (%)
1	40.76	82.36
2	41.38	82.84
3	39.39	80.62
4	38.16	79.70
5	42.56	84.25
6	41.60	83.00
7	41.03	82.74
8	43.05	84.63
9	40.21	82.11
10	40.36	81.84
11	39.39	80.66
12	40.85	82.60
Average	40.73	82.28

The GICam-II compressor has poorer image reconstruction than JPEG and our previous GICam one because the GICam-II compressor uses the subsample scheme to downsample green and blue components according to the

Table 5: The criterion of evaluation.

Score	Description
0~2	Diagnosis is affected
3~5	Diagnosis is slightly affected
6~9	Diagnosis is not affected

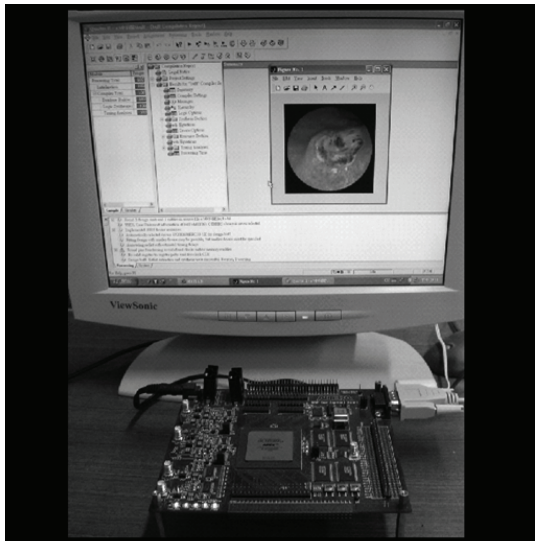


Figure 15: The FPGA prototype of the GICam-II image compressor.

Table 6: The comparison result with previous GICam works.

	JPEG designed by [26]	GICam image compressor [14]	Proposed GICam-II image compressor
Average PSNRY	46.37 dB	41.99 dB	40.73 dB
Average compression rate	82.20%	79.65%	82.28%
Average power dissipation	876 mW	14.92 mW	9.17 mW

2 : 1 and the 4 : 1 subsample ratios. The raw data before compression have lost some raw data information. Hence, the decoded raw data should be reconstructed (the first interpolation) before reconstructing the color images (the second interpolation). Using two-level interpolations to reconstruct the color images has poorer image quality than one-level interpolation. Fortunately, the decoded image quality using GICam-II compressor can be accepted and suitable for practical diagnosis, and the evaluation results of professional gastroenterology doctors can be shown in the last subsection.

Finally, we compare the GICam-II image processor with other works, and the comparison results are shown in Table 7. According to the comparison results, our proposed GICam-II image compressor has lower area and lower operation frequency. It can fit into the existing designs.

Table 7: The comparison results with existing works.

	Area	Frequency (MHz)	Power (mW)	Supply voltage (V)
GICam image compressor [14]	390 k	12.58	14.92 (evaluated)	1.8
Xie et al. [15]*	12600 k	40.0	6.2 (measured)	1.8
Wahid et al. [16]	325 k	150	10 (evaluated)	1.8
Chen et al. [17]*	11200 k	20	1.3 (measured)	0.95
Proposed GICam-II image compressor	318 k	7.96	9.17 (evaluated)	1.8

* Includes analog and transmission circuit and SRAM.

7. Conclusion

In order to further extend the battery life of capsule endoscope, this paper mainly focuses on a series of mathematical statistics to systematically analyze the color sensitivity in GI images from the RGB color space domain to the 2D DCT spatial frequency domain. According to the analysis results, an improved ultralow-power subsample-based GICam image compression processor is proposed for capsule endoscope or swallowable imaging capsules. We make use of the subsample technique to reduce the memory requirements of G1, G2, and B components according to the analysis results of DC/AC coefficients in 2D DCT domain. As shown in the simulation result, the proposed image compressor can efficiently save 38.5% more power consumption than previous GICam one [14] and can efficiently reduce image size by 75% at least for each sampled gastrointestinal image. Therefore, video sequences totally reduce size by 75% at least. Furthermore, the proposed image compressor has lower area and lower operation frequency according to the comparison results. It can fit into the existing designs.

Acknowledgments

This work was supported in part by Chung-Shan Institute of Science and Technology, Taiwan, under the project BV94G10P and the National Science Council, Taiwan, under Grant no. NSC 95-2221-E-009-337-MY3. The authors would like to thank five professional gastroenterology doctors: Dr. Shiann Pan, Dr. Jean-Dean Liu, Dr. Chun-Chao Chang, Dr. Jen-Juh Wang, and Dr. Lou-Horng Yuan from the Division of Gastroenterology, Taipei Medical University Hospital for practical diagnosis, and National Chip Implementation Center (CIC) for technical support. Furthermore, The authors would like to thank Dr. Ping-Kuo Weng and Mr. Yin-Yi Wu because they both design a novel block-based raw image sensor based on the structure of locally raster-scanning.

References

[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, no. 6785, pp. 417–418, 2000.

[2] S. Itoh, S. Kawahito, T. Akahori, and S. Terakawa, "Design and implementation of a one-chip wireless camera device for a capsule endoscope," in *Sensors and Camera Systems for Scientific and Industrial Applications VI*, vol. 5677 of *Proceedings of SPIE*, pp. 109–118, San Jose, Calif, USA, January 2005.

[3] F. Gong, P. Swain, and T. Mills, "Wireless endoscopy," *Gastrointestinal Endoscopy*, vol. 51, no. 6, pp. 725–729, 2000.

[4] H. J. Park, H. W. Nam, B. S. Song et al., "Design of bi-directional and multi-channel miniaturized telemetry module for wireless endoscopy," in *Proceedings of the 2nd Annual International IEEE-EMB Special Topic Conference on Microtechnologies in Medicine & Biology*, pp. 273–276, Madison, Wish, USA, May 2002.

[5] http://www.givenimaging.com/en-int/Patients/Pages/CapsuleEndoscopy.aspx.

[6] http://www.rfsystemlab.com/.

[7] M. Sendoh, K. Ishiyama, and K. I. Arai, "Fabrication of magnetic actuator for use in a capsule endoscope," *IEEE Transactions on Magnetics*, vol. 39, no. 5, pp. 3232–3234, 2003.

[8] L. Phee, D. Accoto, A. Menciassi, C. Stefanini, M. C. Carrozza, and P. Dario, "Analysis and development of locomotion devices for the gastrointestinal tract," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 6, pp. 613–616, 2002.

[9] G. S. Lien, C. W. Liu, M. T. Teng, and Y. M. Huang, "Integration of two optical image modules and locomotion functions in capsule endoscope applications," in *Proceedings of the 13th IEEE International Symposium on Consumer Electronics (ISCE '09)*, pp. 828–829, May 2009.

[10] M. Li, C. Hu, S. Song, H. Dai, and M. Q.-H. Meng, "Detection of weak magnetic signal for magnetic localization and orientation in capsule endoscope," in *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL '09)*, pp. 900–905, Shenyang, China, August 2009.

[11] J. Hou, Y. Zhu, L. Zhang et al., "Design and implementation of a high resolution localization system for in-vivo capsule endoscopy," in *Proceedings of the 8th IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC '09)*, pp. 209–214, Chengdu, China, December 2009.

[12] S.-G. Miaou, S.-T. Chen, and F.-S. Ke, "Capsule endoscopy image coding using wavelet-based adaptive vector quantization without codebook training," in *Proceedings of the 3rd International Conference on Information Technology and Applications (ICITA '05)*, vol. 2, pp. 634–640, July 2005.

[13] S.-G. Miaou, S.-T. Chen, and C.-H. Hsiao, "A wavelet-based compression method with fast quality controlling capability for long sequence of capsule endoscopy images," in *Proceedings of IEEE-EURASIP Workshopon on Nonlinear Signal and Image Processing (NSIP '05)*, p. 34, May 2005.

[14] M. C. Lin, L. R. Dung, and P. K. Weng, "An ultra-low-power image compressor for capsule endoscope," *BioMedical Engineering Online*, vol. 5, article 14, 2006.

[15] X. Xie, G. Li, X. Chen, X. Li, and Z. Wang, "A low-power digital IC design inside the wireless endoscopic capsule," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, pp. 2390–2400, 2006.

[16] K. Wahid, S.-B. Ko, and D. Teng, "Efficient hardware implementation of an image compressor for wireless capsule endoscopy applications," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '08)*, pp. 2761–2765, Hong Kong, June 2008.

[17] X. Chen, X. Zhang, L. Zhang et al., "A wireless capsule endoscope system with low-power controlling and processing ASIC," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 3, no. 1, pp. 11–22, 2009.

[18] C. Hu, M. Q. H. Meng, L. Liu, Y. Pan, and Z. Liu, "Image representation and compression for capsule endoscope robot," in *Proceedings of the IEEE International Conference on Information and Automation (ICIA '09)*, pp. 506–511, June 2009.

[19] J. Wu and Y. Li, "Low-complexity video compression for capsule endoscope based on compressed sensing theory," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '09)*, pp. 3727–3730, Minneapolis, Minn, USA, September 2009.

[20] C. Cheng, Z. Liu, C. Hu, and M. Q.-H. Meng, "A novel wireless capsule endoscope with JPEG compression engine," in *Proceedings of IEEE International Conference on Automation and Logistics (ICAL '10)*, pp. 553–558, Hong Kong, August 2010.

[21] M.-C. Lin, L.-R. Dung, and P.-K. Weng, "An improved ultra-low-power subsample based image compressor for capsule endoscope," in *Proceedings of Medical Informatics Symposium in Taiwan (MIST '06)*, 2006.

[22] H. A. Peterson, H. Peng, J. H. Morgan, and W. B. Pennebaker, "Quantization of color image components in the DCT domain," in *SPIE/IS&T Symposium on Electronic Imaging Science and Technology*, vol. 1453 of *Proceedings of SPIE*, pp. 210–222, San Jose, Calif, USA, February-March 1991.

[23] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 337–343, 1977.

[24] R. W. G. Hunt, *Measuring Colour*, Fountain Press, 1998.

[25] H. R. Kang, *Color Technology For Electronic Imaging Devices*, SPIE Optical Engineering Press, 1997.

[26] V. Bhaskaran and K. kon stantinides, *Images and Video Compression Standards : Alogorithms and Architectures*, Kliwer Academic Publishers, Boston, Mass, USA, 2nd edition, 1997.