



Low-Power Low-Area Near-Lossless Image Compressor for Wireless Capsule Endoscopy

Pawel Turcza¹ · Mariusz Duplaga²

Received: 29 November 2021 / Revised: 28 July 2022 / Accepted: 4 August 2022 /

Published online: 27 August 2022

© The Author(s) 2022

Abstract

The paper presents the concept of a low-power, low-area, near-lossless image compressor for resource-constrained devices such as wireless capsule endoscopy (WCE). The compressor directly processes the raw data from the Bayer Color Filter Array (CFA) imager to avoid the high cost of color interpolation. To improve the efficiency of the compressor in terms of energy consumption, silicon area and compression ratio, the main part of the compressor, i.e., the entropy encoder, uses the existing correlations between the color components of a captured CFA image. The proposed image compressor requires only 12.4% of the memory needed by other high-quality CFA compressors based on the JPEG-LS standard. Despite this significant reduction in memory size, the proposed image compressor outperforms other state-of-the-art coding schemes on capsule endoscopy images. At the same time, it offers only slightly lower performance on standard test images. The proposed image compressor has been implemented as an intellectual property (IP) core using two different low-cost CMOS processes. The design, implemented in UMC 180 nm CMOS process, requires a very low silicon area ($534 \times 426 \mu\text{m}^2$) and consumes very low energy ($22 \mu\text{J}$ per a single 512×512 image frame). Even higher energy efficiency ($12 \mu\text{J}$ per the same image frame) has the IP core implemented in the TSMC 130 nm CMOS process. Both of the selected technologies are low-cost and well-suited to implement a radio frequency transmitter and a low-power successive approximation register analog-to-digital converter in addition to the compressor to provide a cost-effective System on Chip for resource-constrained devices like WCE or wireless camera sensor network.

✉ Pawel Turcza
turcza@agh.edu.pl

Mariusz Duplaga
mariusz.duplaga@uj.edu.pl

¹ AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Krakow, Poland

² Department of Health Promotion and e-Health, Faculty of Health Sciences, Institute of Public Health, Jagiellonian University Medical College, Skawinska Str. 8, 31-066 Krakow, Poland

Keywords Wireless capsule endoscopy · Wireless vision sensor networks · Image compression · Color filter array · Very-large-scale integration (VLSI)

1 Introduction

A wireless capsule endoscope (WCE) is a pill-sized ingestible electronic device incorporating a tiny CMOS image camera with LED lighting, a video processing and control unit, a radio transmitter and batteries. It was developed by Given Imaging Inc. [20] and approved by the US Food and Drug Administration (FDA) for noninvasive examination of the small bowel in adults in August 2001. Nowadays, this revolutionary diagnostic tool is a widely accepted technology that allows a physician to examine the entire gastrointestinal (GI) tract, including the esophagus and the colon [31, 37].

The endoscopic capsule is usually powered by two button-type batteries, which greatly reduces possible energy consumption and takes up space that could otherwise be used to integrate advanced diagnostic and therapeutic functions on board, such as a biopsy mechanism [49] or a precise local drug delivery system [19].

The power consumption of the most popular capsule—the Pillcam SB2—during a standard examination at 4 images per second has been measured and analyzed [2]. It was found that the radio transmitter was the most power-hungry component. The high power consumption of the radio frequency (RF) transmitter is related to the spectral and power restrictions of the medical implant communication service (MICS) band of 401–406 MHz. The permitted channel bandwidth for the MICS band is 300 kHz [51], which significantly limits the rate of data transmission [46] which in turn results in lower picture resolution and lower frame rate when compared with traditional wired endoscopy.

It has been found [52] that random movements of the capsule due to peristalsis, together with a low image quality, were responsible for the overall low level of positive diagnostic results, estimated at under 50%. Consequently, the WCE may miss a significant number of lesions that would be detected by enteroscopy. These limitations have been partially addressed by various companies and have been the focus of intense research by many groups [22, 25, 35, 47]. Currently, the second-generation of the PillCam ESO2 can capture up to 18 fps and has a wider angle of view and automatic light control. However, the Pillcam ESO3 is equipped with two cameras and can record a maximum of 35 fps, which ensures a high detection rate of suspected Barrett’s esophagus and esophagitis [16].

The new PillCam COLON2 has been designed to cope with rapid movements of the capsule, as depending on the measurement of movement, it can change the frame rate from 4 to 35 fps. This approach was improved in [50] and [17] where the capsule was equipped with four and six MEMS-controlled cameras, respectively. To reduce the total power consumption of the capsule and increase the frame rate, the authors chose to store the image data inside on-board flash memory instead of transferring it wirelessly. However, when analyzing the power consumption, the authors [17] found that the built-in flash memory accounted for about 40% of the total energy consumption. It was concluded that applying image compression before storing the image in the flash memory can significantly reduce the overall power consumption of the capsule. To

accommodate the very large amount of data generated by six cameras, the authors apply lossy compression.

Flash memory, instead of the RF transmitter, is also used by Capsocam [40] to reduce power consumption. The Capsocam uses four lateral cameras to capture high-resolution 360° images of the small bowel to facilitate a detailed examination of the mucosal surface. However, the higher image acquisition rate requires a much larger flash memory, which consumes more energy and thus reduces the possible operating time of the capsule in its high frame rate regime. Moreover, the lack of communication prevents the implementation of more advanced functions [24] such as movement, biopsy, or precision local drug release.

It follows that the main approach to solving the transmission bottleneck and reducing the power consumption of the capsule and the space occupied by the batteries should be based on a high-quality, energy-efficient, low-silicon area image compressor [6, 8, 12, 18, 26, 41–43, 47].

WCE captures color images of GI tract using a single-sensor color camera with monochrome CMOS image sensors covered with Bayer color filter array (CFA) [4]. The captured image is composed of 2×2 repeating blocks with two green (G), one red (R), and one blue (B) pixels (Fig. 1a). Based on the acquired CFA data, a full-color image can be obtained by color interpolation [7, 14, 30]. However high-quality color interpolation requires many arithmetic operations on pixel values in a 5×5 window around each recovered pixel [30], requiring a buffer for at least 4 consecutive image lines. Moreover, color interpolation triples the amount of data to be compressed. Therefore, in resource-constrained applications such as capsule endoscopy [1, 6, 8, 18, 41–43] or wireless vision sensor networks [9, 21, 36], in which the overall power consumption and silicon area have to be minimized, compression of a raw CFA data is preferred, while the color interpolation is done using high quality, detail-preserving, demosaicing algorithms in the workstation after data reception.

The inter-pixel correlation in the CFA image is lower compared to the full-color image, which makes well-known techniques such as color transformation [33], predictive or transform coding [17, 28, 43] less efficient for CFA image compression. For these reasons, raw CFA image compression is a very active research topic [1, 6, 8, 27, 42].

Existing image compression methods are generally classified as either reversible (“lossless”) or irreversible (“lossy”). Lossy methods offer higher compression ratios by discarding some information, resulting in inexact image recovery. A near-lossless compression offers a compromise between lossy and lossless compression. In these methods, either the peak of absolute reconstruction error is defined by the user or the peak signal-to-noise ratio (PSNR) between original and reconstructed images is greater than 45–50 dB [5].

Of the recently proposed techniques, the high-performance, near-lossless methods are based on predictive coding [12, 23, 28, 29], the JPEG-LS standard [6, 8, 27] or combine the discrete cosine transform (DCT) with predictive coding [42]. However, the most efficient near-lossless methods [6, 8, 27, 42] need a sophisticated entropy encoder, which requires a large amount of memory for statistical modeling. In this paper, we show how to efficiently exploit existing correlations among the color components in CFA data to significantly reduce the memory requirements of the compressor.

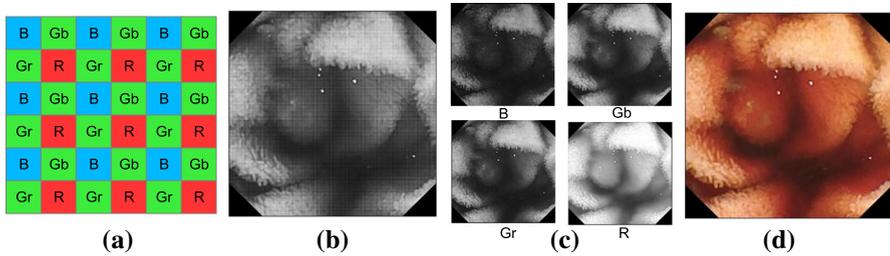


Fig. 1 A sample image in capsule endoscopy. **a** CFA Bayer pattern, **b** CFA mosaic image, **c** Sub-images generated by separating the color components, **d** Full RGB image obtained from the CFA one by color interpolation.

The proposed image compressor uses less than 27% of the memory needed by the previous design [42] and only 12.4% of that needed by the CFA compressor based on JPEG-LS [6, 27] (for an image with 512 columns). The lower memory requirements result in a lower silicon area and greater energy efficiency. Despite the significant resource reduction, the proposed compressor outperforms other state-of-the-art coding schemes on endoscopic images, while it offers only slightly lower performance on standard test images.

The paper is organized as follows. Section 2 analyzes the properties of the endoscopic images processed by the proposed algorithm. The image compressor is presented in detail in Sect. 3. The algorithm's performance in terms of memory requirements and the PSNR versus bit-rate in bits-per-pixel (bpp) is analyzed and compared with other related compressors in Sect. 4. Hardware implementation aspects and the results of the VLSI implementation are presented in Sect. 5. Section 6 contains the conclusions.

2 Analysis of the Properties of Endoscopic Images

The design of an efficient image compression algorithm for wireless capsule endoscopy should be preceded by an analysis of the image properties of the human gastrointestinal system. Such an analysis, for full-color images, can be found in the works [12, 23, 28, 29]. In particular, it was found that the correlation of the color components of RGB color space is very high. In addition, the cross-correlation between the green and blue channels is the highest in endoscopic images. In contrast, in natural images, the cross-correlation is the highest between the red and green channels. Based on this observation efficient but simple algorithms based on differential pulse code modulation (DPCM) of decorrelated color planes were proposed [12, 23, 28, 29]. The resulting prediction residuals were encoded using separate entropy encoders.

A similar approach was taken in work [8] but for coding WCE images directly in CFA color space. However, in a CFA image, each pixel has only one color component Fig. 1. To perform the color transformation of CFA data, one requires pixel values from four different locations, which span two consecutive image lines. This, in turn, significantly increases the cost of color transformation of CFA data [7, 14]. Moreover,

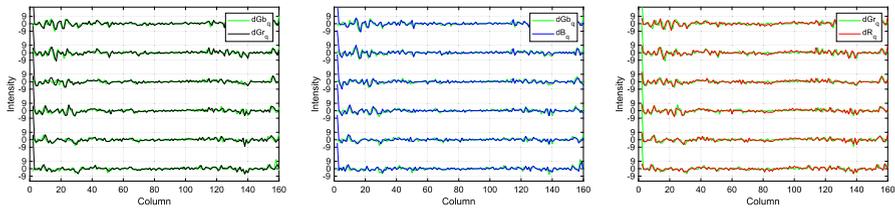


Fig. 2 Line plots showing the similarity between the quantized prediction residuals of color channels along 6 rows taken from the central part of sub-images in Fig. 1c. dBq_q and $dGrq_q$ denote quantized prediction residuals of green samples from even and odd lines, respectively (see Fig. 1a).

the color transformation matrix is usually not orthogonal, which leads to an increase in the dynamic range of color transformed components and an increase in quantization noise at the decoder [39]. Although this is not a significant issue for lossy algorithms, it can be a disadvantage for near-lossless methods, especially for CFA images. In addition, low inter-pixel correlation in CFA images, compared to full-color ones, makes color transformation less efficient in such cases.

Although the inter-pixel correlation in CFA images is lower than in full-color images, it is still high enough to make differential coding a useful tool for compressing such images. Moreover, it can be seen (see Fig. 2) that there is quite a significant cross-correlation between the pixels of the color components encoded by DPCM. This pixel-domain observation can be confirmed by analyzing the normalized cross-correlation (NCC) among different color components encoded with DPCM (see Fig. 3):

$$\phi_{dX,dY}(k) = \frac{\sum_{r,c} dX(r,c)dY(r,c-k)}{\sqrt{\sum_{r,c} dX^2(r,c)}\sqrt{\sum_{r,c} dY^2(r,c)}}, \tag{1}$$

where $dX(r,c)$, and $dY(r,c)$ represent the values of the prediction residuals of two color components X , and Y of CFA image at the point (r,c) , and k is a row-wise offset.

Based on the above observation, we offer an image compression scheme that skips the color space transformation step and performs the DPCM coding directly on the individual color channels of the CFA image. The remaining similarity between prediction residuals of different color components is used by the entropy encoding step to improve the performance of the encoder in terms of the size of the required memory and the compression ratio.

3 Algorithm Description

The block diagram of the proposed near-lossless image compression algorithm is shown in Fig. 4, and its data processing steps are shown in Fig. 5. The proposed image compression scheme omits the color space transformation and applies the DPCM predictive coding to each color component separately in the first step. The resulting prediction residuals dBq , dGq , and dRq are entropy encoded using the single context-adaptive Golomb-Rice encoder [34]. To improve the efficiency of the entropy encoder

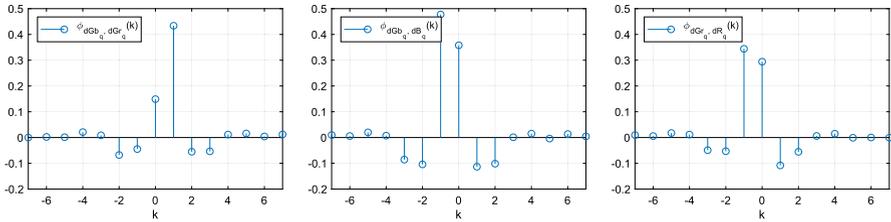


Fig. 3 Normalized cross-correlation (1) value among the quantized prediction residuals of different color components of the sample CFA image from Fig. 1c.

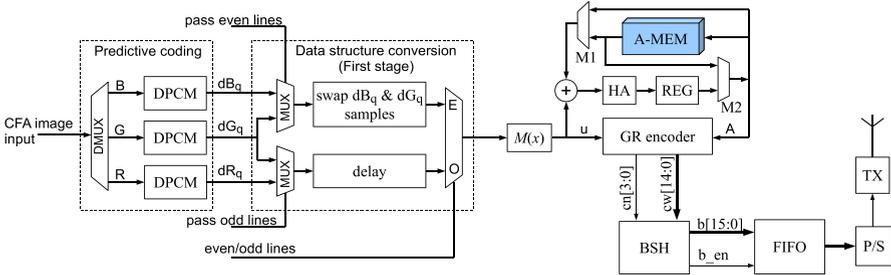


Fig. 4 Block diagram of the proposed near-lossless image compression algorithm.

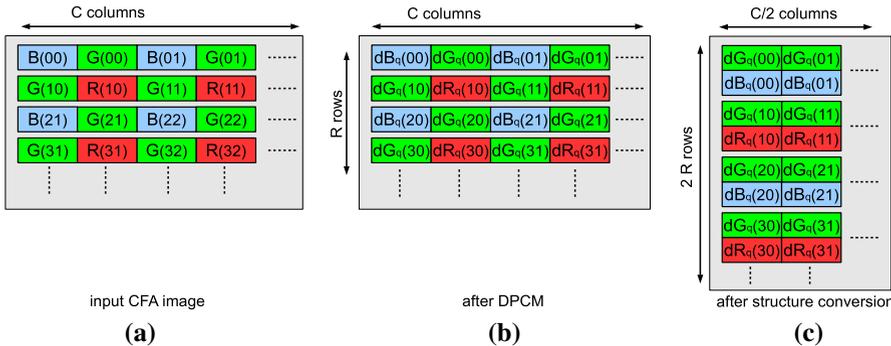


Fig. 5 CFA data processing steps.

in terms of compression ratio and memory requirements, the data structure conversion right after DPCM is performed.

3.1 Predictive Encoding

Endoscopic images are smooth, so sharp edges are rare. Therefore, differential pulse code modulation can be particularly efficient in removing spatial redundancy in endoscopic images. In image coding application, the prediction error $dX(r, c)$ in the

conventional DPCM (see Fig. 6a) is calculated as

$$dX(r, c) = X(r, c) - X_p, \quad (2)$$

where $X(r, c)$ represents the actual pixel value, while X_p is its prediction from previously coded pixels (known at the coder and decoder). In JPEG-LS, prediction depends on three already encoded adjacent pixels, which together with a sophisticated entropy encoder, makes JPEG-LS a very efficient scheme. However, JPEG-LS requires a significant amount of memory in form of two blocks. The first block is used to store the previous image line for the predictor, and the second block consists of 365×34 bits of context variables for entropy encoding. The high memory requirement results in a large silicon area and high power consumption. Therefore, the proposed algorithm assumes a simpler, one-dimensional predictor. For the conventional DPCM (see Fig. 6a) and the green samples the predictor operates according to the formula:

$$X_p = \begin{cases} \hat{X}(r, c - 1), & c > 0 \\ \left\lfloor \left(\hat{X}(r - 1, 0) + \hat{X}(r - 1, 1) \right) / 2 \right\rfloor, & c = 0, r > 0. \\ 0, & c = 0, r = 0 \end{cases} \quad (3)$$

For red and blue samples, the slightly simpler predictor is used:

$$X_p = \begin{cases} \hat{X}(r, c - 1), & c > 0 \\ \hat{X}(r - 1, 0), & c = 0, r > 0. \\ 0, & c = 0, r = 0 \end{cases} \quad (4)$$

To increase the coding efficiency of full-color endoscopic images, it was proposed in [12] to precede the prediction encoding with a scalar quantizer (see Fig. 6b). Therefore, the modified DPCM (Fig. 6b) uses a predictor that works like (3) and (4) but on quantized image samples, i.e., $\hat{X}_q(r, c)$.

In this paper, in application to direct compression of CFA endoscopic images, we evaluate the efficiency of the two versions of DPCM shown in Fig. 6. Both versions use a uniform scalar quantizer Q . The quantization and its inverse are defined as:

$$x_q = Q(x) = \text{round}(x/\delta), \quad (5)$$

$$\hat{x} = Q^{-1}(x_q) = \delta \cdot x_q, \quad (6)$$

where δ is the quantization factor selected by the user. Quantization is an irreversible process. The proposed compressor is strictly near-lossless. For $\delta = 4$, the reconstruction error for any pixel does not exceed the value of 2 intensity levels, which corresponds to the parameter $NEAR = 2$ in JPEG-LS.

The only difference between the two considered DPCM schemes is the placement of the quantizer (5). The numerical results presented in Sect. 4 show that the scheme (b) offers better performance.

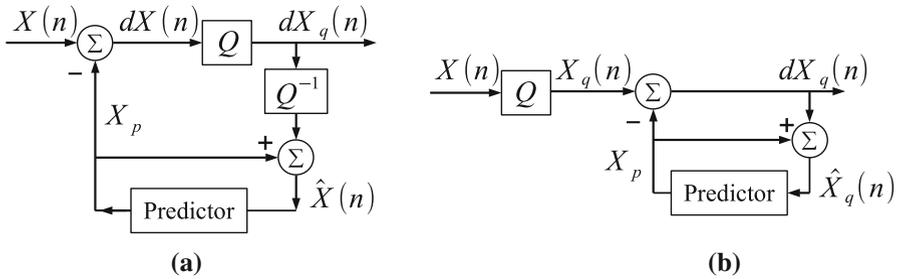


Fig. 6 Two versions of DPCM encoders, **a** with the embedded quantizer, and **b** the quantizer first.

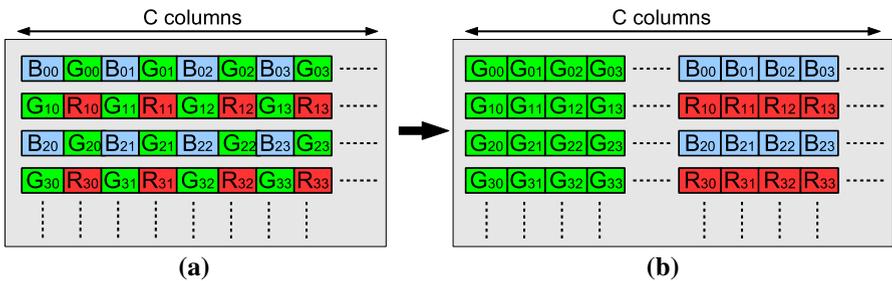


Fig. 7 Structure conversion of the CFA image in the compression scheme [6] based on JPEG-LS.

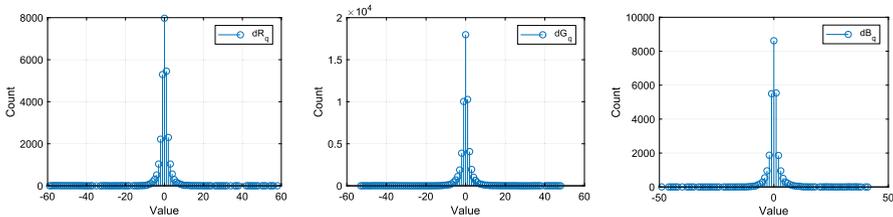
3.2 Data Structure Conversion

In the previous capsule endoscopy image compression scheme [6] based on JPEG-LS, the structure conversion was applied to group the *G* components on the left side of the frame, while the *R* and *B* components were shifted to the right of the frame (see Fig. 7). This operation allows using of a single JPEG-LS engine for coding all three color components in the CFA image. However, the overall memory requirements of such a modified codec [6] remain still very high (see Table 1).

Therefore in this paper, based on the observed correlation properties (see Fig. 3), we propose a structure conversion, which involves two simple operations. Firstly, the prediction residual samples on the even lines, i.e., dB_q and dG_q , are swapped, then the samples of green channel prediction residuals (dG_q) are put into even lines, while the samples of prediction residuals from red or blue channels, i.e., dR_q or dB_q constitute the odd lines. The resulting image is presented in Fig. 5c. The purpose of the structure conversion in the proposed algorithm is to establish the context for the entropy encoder. For each column in the image in Fig. 5c a separate context is stored in A-MEM. It should be stressed that the proposed structure conversion takes place on-the-fly during the entropy encoding stage. Therefore, it does not require a large block storage area, just a single sample in addition to A-MEM. The proposed structure conversion allows using a single context adaptive Golomb Rice entropy encoder to compress prediction residuals from all three DPCM channels, which in turn reduces coder memory requirements 3 times as it is shown in the next section.

Table 1 Memory (SRAM) requirements for different image compressors (for an image with 512 columns).

Memory	JPEG-LS [6, 27]	Work [8]	Work [1]	Work [42]	Proposed
Predictor	512×8	1024×8	256×8	192×8	0
GR encoder	365×34	0	0	768×8	256×8
Total [bits]	16506	8192	2048	7680	2048

**Fig. 8** Histograms of quantized prediction residuals of R , G and B CFA channels of image from Fig. 1.

3.3 Context Adaptive Golomb-Rice Encoder

Entropy encoding is a lossless coding technique that is applied to the prediction residuals to achieve further data compression. It involves two separate and independent operations: modeling and coding. Typically, the model is determined based on the data statistics. Figure 8 shows that prediction residuals dR_q , dB_q and dG_q can be modeled by a two-sided geometric (TSG) distribution centered at zero [45]. Moreover, for an 8-bit CFA image and the quantizer factor $\delta \geq 4$, prediction residuals can be limited to the range $[-63, 63]$.

The entropy encoder memory requirements depend primarily on the number of contexts used during coding. Therefore, to reduce the encoder memory size, the correlation properties of prediction residuals were analyzed (see Fig. 3). It was observed that the prediction residuals from different color channels show high similarity. Figure 2 shows that the cross-correlation among prediction residuals taken from neighboring image lines of the green channel is nearly as high as the cross-correlation among prediction residuals of samples from the same image line but different channels. The existing similarity can be used to reduce the number of contexts. Instead of using a separate context for each of the columns of each color component as in the work [42], a common context is used to encode all prediction residuals from all of the channels in the same column. The proposed approach reduces the encoder memory size three times compared to the work [42], i.e., from $3 \times C/2$ down to $C/2$, where C is the number of columns in the CFA image. The proposed entropy encoder requires 6 times less memory (see Table 1) compared to the encoder from the JPEG-LS. The numerical results presented in the next section show that despite significant memory reduction, the performance of the proposed algorithm remains very competitive with significantly more complex JPEG-LS, especially for coding capsule endoscopy images.

To meet low-complexity requirements and take full advantage of the TSG distribution the proposed encoder uses Golomb-Rice (GR) codes. The GR coder was designed

to encode sequences of non-negative numbers, so the prediction residuals dR_q, dG_q, dB_q , before encoding, are mapped to non-negative integers in a reversible manner [34]:

$$u = M(x) = \begin{cases} 2x, & x \geq 0 \\ 2|x| - 1, & x < 0 \end{cases}. \quad (7)$$

A Golomb-Rice code is a variable-length code. The encoded integer u is split into two parts: the quotient $q = \lfloor u/2^k \rfloor$ and the remainder $r = u - 2^k q$, which depend on a single tunable parameter $k \geq 0$. The quotient is sent in unary coding using $q + 1$ bits and the remainder is encoded in binary using k bits. The length $cn = \lfloor u/2^k \rfloor + 1 + k$ of the generated codeword cw (see Fig. 4) depends on the coded integer $0 \leq u \leq 127$ and the parameter k . Therefore, the Golomb-Rice encoder allows for a clear separation of modeling and coding operations. In the proposed compressor, the parameter k is tuned symbol-by-symbol using formula [42], [32]:

$$k = \min \left\{ i : 2^i \cdot N > A, i \geq 0 \right\}. \quad (8)$$

In (8), A represents the sum of magnitudes of encoded prediction residuals (2), while N is their number. In order to obtain the highest possible compression, the parameter k is adapted to local statistics of the coded data which vary in the image plane. In the proposed coder, the parameter k is tuned separately for each column of the encoded block (Fig. 5c). To speed up the adaptation, the values of the registers N and A are halved (using the module HA in Fig. 4) each time N equals the specified threshold N_0 . It has been found that the optimal threshold for the proposed encoder is $N_0 = 4$. To limit the computational cost of (8), the values k are confined to the range $0 \leq k \leq 4$.

The resulting variable-length codewords cw (see Fig. 4) are aligned and packed into 16-bits words by the barrel shifter (BSH). The actual length of the codeword is available at the output cn . Suppose the length of unary part $q + 1$ of the codeword cw is greater than 7, which happens when the parameter k is chosen inappropriately. In that case, the GR encoder is bypassed to prevent overflow of the BSH. In that case, the encoded integer u is represented by an escape sequence equal to 255 followed by u in 7 bits.

The rate of the 16-bit words stream at the BSH output depends on the amount of fine detail in the compressed image. Since the transmitter (TX) operates with a constant rate, a FIFO is used to average the data rate. The parallel-to-serial (P/S) converter on the FIFO output is used to serialize the data to drive the RF transmitter.

4 Algorithm Performance

Memory size and its access frequency strongly influence the silicon area and power consumption of the image compressor. The proposed compressor (see Fig. 4) needs only one storage area (A-MEM). The A-MEM block stores values of A for (8) separately for each column of the encoded data block (see Fig. 5c), and therefore needs $C/2$ words of 8 bits. The memory requirements for different compressors, for images with

$C = 512$ columns, are given in Table 1. The proposed algorithm assumes a simple, 1D horizontal prediction (3) and (4) so the memory is required only by the GR encoder (A-MEM). It can be seen that the entropy encoder in the new compressor requires only $1/3$ of the storage area (A-MEM) of the previous version [42]. The whole image compressor requires less than 27% of the memory required by the original [42] and only 12.4% of the memory required by the compressors [6, 27] based on the JPEG-LS standard.

Figure 5c shows that coefficients from different color channels are encoded alternately, and the green channel is always encoded first. Therefore, the values of A_j , stored in A-MEM, after encoding the green channel prediction residual sample $dG_q(i, j)$ can be buffered, using the REG register (see Fig. 4) and multiplexers (M1 and M2), and reused for encoding samples of $dB_q(i, j)$ and $dR_q(i, j)$ representing respectively the prediction residuals of blue and red components of the CFA image. Such an operation reduces the access frequency of the A-MEM memory by 50%, decreasing the power consumption.

The performance, in terms of the $PSNR$ versus the required bit-rate in bpp , of the proposed image compressor is evaluated on the basis of images from four different datasets. The $PSNR$ was computed in the CFA image space using the following equation:

$$PSNR(dB) = 10 \log_{10} \frac{255^2}{MSE}, \quad (9)$$

where $MSE = \langle (x_i - \hat{x}_i)^2 \rangle$ is the mean square error, $\langle \cdot \rangle$ denotes averaging operation, and x_i and \hat{x}_i are the values of the pixels in the original and reconstructed CFA image, respectively.

Table 2 shows the results for the standard test images and includes comparisons with other state-of-the-art methods. Tables 3, 4, and 5 present the results for medical images obtained from wired endoscopy and capsule endoscopy. The wired endoscopy images were provided by partners in the VECTOR [35] project and were used to evaluate the performance of our previous WCE image compressors [41–43]. The capsule-endoscopy images dataset includes images from the PillCam and the Olympus EC-S10 capsules. The PillCam dataset images were extracted from videos provided by Given-Imaging and Clinica CEDIG on their official YouTube channels [10, 15]. The Olympus EC-S10 dataset consists of sample images (Fig. 9) from various labeled classes in the Kvasir-Capsule dataset [38]. Images from capsule-endoscopy have blacked-out corners, so those areas are skipped by the compressor to save the bits [23], [42].

Table 2 shows that for standard test images, the proposed algorithm offers slightly lower performance than the JPEG-LS-based algorithms [6, 26] and the previous DCT-based encoder [42]. In particular, it can be seen that while the $PSNR$ is comparative, the required bit-rate is about 6% higher. The situation improves gradually for medical images. From Table 3, it can be seen that the proposed scheme offers nearly the same image quality as JPEG-LS-based coders [6] and [8], developed especially for compression of endoscopic images, but achieves it at a much lower bit-rate. Table 4 shows the results for images from PillCam. It can be observed that the proposed scheme offers the highest $PSNR$ at a negligible higher bit-rate.

Table 2 Bitrate (bpp) and PSNR (dB) for standard test images for different coders.

Image	Airplane	Baboon	House	Lenna	Peppers	Average	
	bpp	bpp	bpp	bpp	bpp	bpp	PSNR
JPEG-LS	3.012	5.159	3.933	4.895	5.236	4.447	45.16
Work [6]	2.983	4.912	3.474	3.464	3.527	3.672	46.43
Work [26]	2.971	4.651	3.227	3.408	3.423	3.536	46.30
Work [8]	3.280	5.397	3.839	3.838	3.821	4.035	46.42
Work [42]	2.937	4.633	3.454	3.229	3.311	3.513	46.43
Work [1]	1.507	1.776	1.409	1.521	1.511	1.545	30.62
DPCM (a)	3.168	4.974	3.603	3.688	3.683	3.823	46.44
DPCM (b)	3.000	4.947	3.470	3.594	3.602	3.723	46.46

Table 3 Bitrate and PSNR (dB) for push-endoscopy images [42].

	Image	(a)	(b)	(c)	(d)	(e)	(f)	Average
JPEG-LS	bpp	3.901	3.566	3.562	3.935	2.532	2.900	3.399
	PSNR	45.09	45.08	45.13	45.11	45.12	45.13	45.11
Work [6]	bpp	3.457	3.111	3.422	2.894	3.935	2.532	3.225
	PSNR	46.42	46.37	46.42	46.39	46.41	46.44	46.41
Work [8]	bpp	3.813	3.531	3.993	3.207	2.658	2.509	3.285
	PSNR	46.37	46.36	46.38	46.39	46.37	46.38	46.37
Work [1]	bpp	1.522	1.361	1.482	1.305	1.144	1.065	1.313
	PSNR	34.65	35.85	37.23	38.10	38.48	38.95	37.21
Work [42]	bpp	3.270	2.915	3.308	2.803	2.396	2.251	2.824
	PSNR	46.53	46.52	46.51	46.53	46.57	46.61	46.55
DPCM (a)	bpp	3.534	3.239	3.711	2.970	2.463	2.401	3.053
	PSNR	46.36	46.35	46.39	46.36	46.37	46.41	46.37
DPCM (b)	bpp	3.426	3.104	3.632	2.837	2.298	2.238	2.923
	PSNR	46.32	46.29	46.38	46.38	46.38	46.42	46.361

The results for the Olympus EC-S10 capsule are presented in Table 5. It can be seen that the new compressor, despite very low memory requirements (see Table 1), outperforms JPEG-LS-based coders [6] and [8] in terms of $PSNR$ and bpp . Compared to the previous DCT-based codec [42], it offers a slightly lower $PSNR$, but at the lower required bit-rate. The lowest bit-rate is offered by [1] but at a much lower $PSNR$. The significant reduction in bpp in [1] is achieved by strong quantization and down sampling of the green channel of the CFA source image, which induces significant color distortion in richly detailed images.

Table 4 Bitrate (bpp) and PSNR (dB) for PillCam images [42].

	Image	(a)	(b)	(c)	(d)	(e)	(f)	Average
JPEG-LS	bpp	3.157	2.818	3.072	3.050	2.703	2.531	2.870
	PSNR	45.37	45.25	45.27	45.48	45.47	45.41	45.38
Work [6]	bpp	2.252	2.164	1.816	2.086	2.216	1.904	2.060
	PSNR	46.41	46.41	46.45	46.44	46.43	46.42	46.43
Work [8]	bpp	2.453	2.385	1.795	2.304	2.534	2.122	2.266
	PSNR	46.41	46.39	46.31	46.43	46.46	46.44	46.41
Work [1]	bpp	1.132	1.118	1.004	1.118	1.159	1.040	1.095
	PSNR	39.29	39.51	39.55	39.35	39.02	40.01	39.45
Work [42]	bpp	2.210	2.139	1.766	2.120	2.253	1.919	2.068
	PSNR	46.55	46.53	46.92	46.59	46.50	46.66	46.62
DPCM (a)	bpp	2.325	2.290	1.993	2.256	2.380	2.151	2.232
	PSNR	46.86	46.83	46.75	46.80	46.75	46.84	46.805
DPCM (b)	bpp	2.217	2.174	1.845	2.139	2.283	2.031	2.115
	PSNR	46.87	46.83	46.81	46.79	46.84	46.85	46.830

The efficiency of the entropy encoder can also be assessed by comparing the resulting bitrate in *bpp* to the Shannon entropy

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i), \quad (10)$$

where $p(x_i)$ is the probability of the symbol x_i from alphabet $X = \{x_1, x_2, \dots, x_n\}$.

The proposed encoder uses a separate GR parameter (8) to encode each of the columns of the data block. Therefore, to ensure a fair comparison of the entropy measure, the average value of the Shannon entropy (10) computed separately for each column and each color plane of the encoded data block was used. The results are presented in Table 6. It can be seen that the proposed encoder achieves a bit-rate very close to the Shannon entropy, which confirms the efficacy of the new encoder in adaptation to local statistics of encoded data.

5 VLSI Implementation

The offered image quality in relation to the compression ratio, the required power consumption, and the silicon area are the key parameters in the design and evaluation of resource-constrained devices such as WCE. The image quality over compression ratio has been successfully assessed in the previous section using MATLAB. To evaluate the required silicon area and the related power consumption the proposed algorithm was implemented using two different CMOS processes. The proposed image compressor together with additional blocks: camera interface, bit stream FIFO (1 KB), and data serializer [42] was implemented in the Verilog Hardware Description Language and

Table 5 Bitrate (bpp) and PSNR (dB) for Olympus EC-S10 images [38].

	Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	Average
JPEG-LS	bpp	2.610	3.949	2.736	3.187	3.557	2.754	3.576	3.601	3.945	3.509	2.555	3.279	3.173	3.264
	PSNR	45.31	45.30	45.32	45.33	45.33	45.20	45.29	45.35	45.27	45.31	45.31	45.30	45.31	45.30
Work [6]	bpp	1.942	2.770	2.068	2.022	2.781	2.191	2.423	1.923	2.110	2.152	2.129	2.932	2.159	2.277
	PSNR	46.44	46.47	46.48	46.45	46.45	46.46	46.43	46.41	46.45	46.43	46.43	46.44	46.44	46.44
Work [8]	bpp	1.974	2.970	2.186	2.157	3.088	2.355	2.628	1.966	2.298	2.307	2.246	3.216	2.382	2.444
	PSNR	46.57	46.54	46.50	46.52	46.50	46.53	46.51	46.51	46.54	46.53	46.49	46.46	46.51	46.52
Work [1]	bpp	1.040	1.288	1.143	1.129	1.377	1.183	1.268	1.074	1.135	1.111	1.149	1.409	1.187	1.192
	PSNR	32.93	30.77	31.10	33.02	30.94	32.20	31.48	32.26	33.15	34.26	33.05	32.17	32.08	32.26
Work [42]	bpp	1.995	2.755	2.074	2.063	2.798	2.232	2.409	1.951	2.122	2.197	2.162	2.969	2.186	2.301
	PSNR	47.36	46.91	47.20	47.00	46.84	47.06	46.90	47.27	47.02	47.14	47.16	46.82	46.95	47.05
DPCM (a)	bpp	2.063	2.819	2.228	2.148	2.852	2.348	2.458	2.077	2.227	2.281	2.253	2.994	2.296	2.388
	PSNR	46.60	46.66	46.57	46.63	46.60	46.61	46.62	46.62	46.60	46.65	46.59	46.61	46.63	46.614
DPCM (b)	bpp	1.853	2.683	2.020	1.951	2.711	2.158	2.286	1.868	2.024	2.101	2.052	2.866	2.102	2.206
	PSNR	46.59	46.66	46.54	46.63	46.59	46.61	46.64	46.66	46.58	46.65	46.59	46.60	46.63	46.613

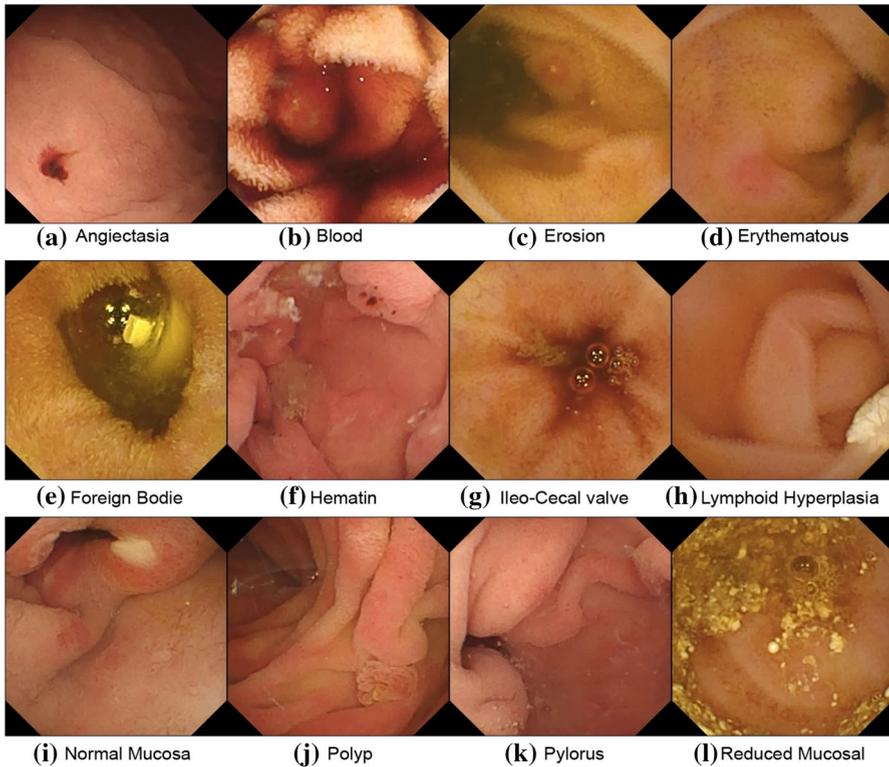


Fig. 9 Olympus EC-S10 test images from Kvasir-Capsule dataset [38] published under CC BY-NC 4.0 license. <http://creativecommons.org/licenses/by/4.0/>.

then synthesized using the Cadence RTL Compiler to UMC 180nm and TSMC 130nm CMOS processes. For the design in the UMC L180 MM/Rf 1P6M process, the standard cell library (FSA0M_A) from Faraday Technology Corporation [13] was used. The required static random access memories (SRAM) were provided to us by EURO-PRACTICE [11] using a memory compiler (FSA0A_C_SJ), also from Faraday. The design using TSMC 130nm CL013G-FSG process was implemented using a 7-tracks standard cell library from the Artisan [3]. The required memories were generated using a memory compiler from Artisan. Both of the selected technologies are low-cost and well-suited to implement a radio frequency transmitter [48] and a low power successive approximation register (SAR) analog-to-digital converter (ADC) [44] in addition to the compressor to provide a cost-effective System on Chip (SoC) for WCE.

The layouts of the two designed IP cores are shown in Fig. 10. It can be seen that the SRAM blocks occupy the largest part of the silicon area. The area of the compressor IP core implemented in the UMC 180nm process is $534 \times 426 \mu\text{m}^2$, which is about 44% of the area of our previous design [42] made with the same technology to allow for a fair comparison. The design implemented in the TSMC 130nm CMOS process requires only $370 \times 356 \mu\text{m}^2$ of silicon.

Table 6 Bitrate (*bpp*) and average entropy of quantized coefficients computed separately for each column of encoded block (*ent*) and for each color plane and each column of encoded block (*ent3*) for Olympus EC-S10 images [38].

Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	Average
DPCM (a)	2.063	2.819	2.228	2.148	2.852	2.348	2.458	2.077	2.227	2.281	2.253	2.994	2.296	2.388
DPCM (a)	2.032	2.936	2.152	2.083	2.942	2.385	2.443	2.028	2.182	2.302	2.264	3.169	2.261	2.399
DPCM (a)	1.991	2.875	2.111	2.043	2.882	2.337	2.391	1.986	2.143	2.256	2.219	3.104	2.221	2.351
DPCM (b)	1.853	2.683	2.020	1.951	2.711	2.158	2.286	1.868	2.024	2.101	2.052	2.866	2.102	2.206
DPCM (b)	1.801	2.817	1.939	1.898	2.825	2.197	2.299	1.792	1.995	2.114	2.056	3.063	2.081	2.221
DPCM (b)	1.760	2.755	1.899	1.859	2.766	2.150	2.248	1.752	1.956	2.070	2.012	2.998	2.042	2.174

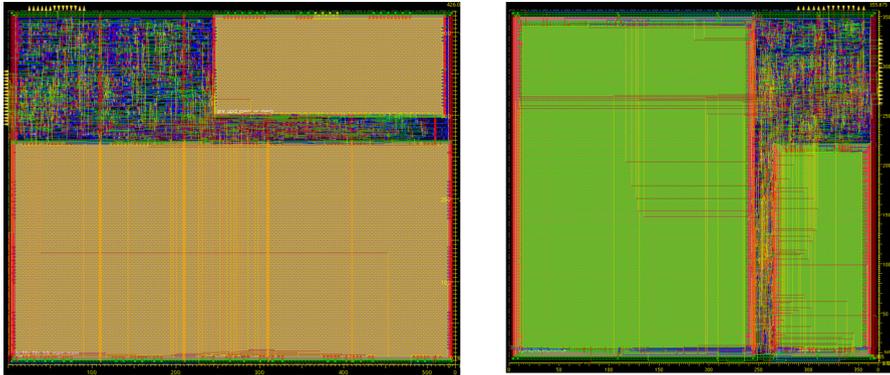


Fig. 10 Layout views of the IP cores for image processing systems for WCE designed in the UMC 180nm (left) and TSMC 130nm (right) technologies.

Table 7 Breakdown of the power consumed by the design operating at 20 fps.

Power	Internal [μ W]	Switching [μ W]	Leakage [μ W]	Total [μ W]	Percentage [%]
Group	Technology UMC 180nm				
Sequential	111.5	8.95	0.84	121.3	27.82
Macro	112.5	0.51	6.52	119.5	27.4
Combinational	67.1	60.73	0.96	128.8	29.52
Clock Tree	26.6	39.99	0.04	66.54	15.26
Total	317.8	110.1	8.35	436.2	100
Group	Technology TSMC 130nm				
Sequential	50.99	1.95	3.103	56.04	23.74
Macro	27.56	0.10	87.2	114.9	48.65
Combinational	22.32	16.62	5.73	44.67	18.92
Clock Tree	9.60	10.70	0.22	3.33	1.41
Total	110.5	29.38	96.25	236.1	100

After the design had been placed and routed, a thorough power analysis was performed. The results obtained with the Cadence (Genus / Innovus) ASIC design tools for a compressor running at 20 fps with a resolution of 512×512 , which requires a clock frequency of 5.25 MHz, are shown in Table 7. The total power consumed by the design has been broken down into internal power that is consumed by all logic cells and macrocells such as SRAM, switching power due to parasitic capacitances, and static power due to leakage current.

A comparison with other designs is shown in Table 8. It can be seen that the new design has very low energy consumption. The design using the UMC 180nm process requires 22μ J per a single 512×512 image frame, which is 50% of the previous design [42]. The energy efficiency of the IP core implemented in the TSMC 130nm process is even higher, as it requires only 12μ J per the same image frame.

Although the work [8] reports lower power consumption, its compression ratio is about 8% lower than the proposed version. In addition, the required line memory

Table 8 Comparison with other ASIC design for WCE (w/o means that FIFO was not included).

Work	[8]	[27]	[42]	[18]	[43]	Proposed
Algorithm	DPCM	JPEG-LS	DCT	2D-DCT	2D-DCT	DPCM
Technology [nm]	180	180	180	180	180	180
Image resolution	640 × 480	400 × 400	512 × 512	480 × 480	320 × 320	512 × 512
Frames/sec	81	24	20	2	50	20
Power [mW]	0.9	4.2	0.88	1.03	1.32	0.44
Energy [μ J/frame]	11.1	190	44	515	27	12
Energy [nJ/pixel]	0.036	1.18	0.167	2.23	0.26	0.046
Memory [kbits]	10.2	100	7.68	36	20,48	2.0
FIFO [kbits]	(off-chip) w/o	(on-chip) (total)	(on-chip) 8,192	(on-chip) (total)	(on-chip) 32,768	(on-chip) 8,192
Logic cells	12.4k	27.8k	1.8k	30k	4.2k	994
						1146

buffer (10.2 kbit) was implemented in the work of [8] as off-chip memory, so its power consumption is unknown and is therefore not included in the chip's figure of merit (Energy [nJ/pixel]). The design [1] was only partially implemented in FPGA, so it is not included in Table 8.

6 Conclusions

In this paper, a low-power, low-resources CFA image compressor for WCE application has been proposed. It has been shown that the memory requirements of the compressor can be greatly reduced through proper data structure conversion before entropy encoding. The proposed data structure conversion reduces the entropy encoder memory size by a factor of 3 and the access frequency by a half. As a result, the total memory required by the compressor has been cut by more than 70% compared to the previous DCT-based codec and by 87% compared to JPEG-LS based codec. It has been shown that the proposed compressor outperforms other state-of-the-art coding schemes on capsule endoscopic images, while it offers only slightly lower performance on standard test images. The high efficiency of the entropy coder alone, the main part of the image compressor, has been confirmed by showing that the average bit-rate expressed in bits per pixel is very close to the Shannon entropy. The silicon area of the designed compressor core has been reduced by 56% compared to the previous design using the same technology node. In addition, the total power consumption has been cut by 50%. The very small silicon area and low energy consumption make the proposed compressor ideal for resource-limited video applications such as wireless capsule endoscopy or wireless vision sensor networks.

Acknowledgements This work was supported by the Polish Ministry of Science and Higher Education with the subvention funds of the Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering of AGH University of Science and Technology, Grant No 16.16.120.773. The authors would like to thank Erwin Deumens ASIC Design support Engineer from EURORACTICE IC service for providing EDA-views of SRAM memory for the design in UMC 180nm technology.

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Q. Al-Shebani, P. Premaratne, P.J. Vial, D.J. McAndrew, The development of a clinically tested visually lossless Image compression system for capsule endoscopy. *Signal Process. Image Commun.* **76**, 135–150 (2019). <https://doi.org/10.1016/j.image.2019.04.008>

2. Q. Angermann et al., Smart videocapsule for early diagnosis of colorectal cancer: toward embedded image analysis, in *Computational Intelligence in Digital and Network Designs and Applications*. (Springer, Cham, 2015), pp.325–350. https://doi.org/10.1007/978-3-319-20071-2_12
3. ARM Artisan Physical IP: <https://developer.arm.com>. Accessed 28 July 2022
4. B. E. Bayer (1976) Color imaging array, US Patent 3:971 065
5. A. Bazhyna, K. Egiazarian, Lossless and near lossless compression of real color filter array data. *IEEE Trans. Consum. Electron* **54**(4), 1492–1500 (2008). <https://doi.org/10.1109/TCE.2008.4711192>
6. X. Chen, X. Zhang, L. Zhang, X. Li, N. Qi, H. Jiang, Z. Wang, A wireless capsule endoscope system with low power controlling and processing ASIC. *IEEE Trans. Biomed. Circ. Syst.* **3**(1), 11–22 (2009). <https://doi.org/10.1109/TBCAS.2008.2006493>
7. S.L. Chen, H.R. Chang, Fully pipelined low-cost and high-quality color demosaicking VLSI design for real-time video applications. *IEEE Trans. Circuits Syst. II, Express Briefs* **62**(6), 588–592 (2015). <https://doi.org/10.1109/TCSII.2015.2407773>
8. S.L. Chen, T.Y. Liu, C.W. Shen, M.C. Tuan, VLSI implementation of a cost-efficient near-lossless CFA image compressor for wireless capsule endoscopy. *IEEE Access*. **4**, 10235–10245 (2016). <https://doi.org/10.1109/ACCESS.2016.2638475>
9. S.L. Chen, J. Nie, T.L. Lin, R.L. Chung, C.H. Hsia, T.Y. Liu, H.X. Wu, VLSI implementation of an ultra-low-cost and low-power image compressor for wireless camera networks. *J. Real-Time Image Process.* **14**(4), 803–812 (2018). <https://doi.org/10.1007/s11554-015-0553-z>
10. Clinica CEDIG YouTube channel: <https://www.youtube.com/watch?v=4uiefWnZBRk>. Accessed 28 July 2022
11. EURO PRACTICE IC SERVICE: <https://europractice-ic.com>. Accessed 28 July 2022
12. K.A. Fante, B. Bhaumik, S. Chatterjee, Design and implementation of computationally efficient image compressor for wireless capsule endoscopy. *Circ. Syst. Sig. Proces.* **35**(5), 1677–1703 (2016). <https://doi.org/10.1007/s00034-015-0136-z>
13. Faraday silicon IP: <https://www.faraday-tech.com/en/entry/SiliconIP>. Accessed 28 July 2022
14. B. Garg, S.K. Rai, A. Puranik, G.K. Sharma, ES-COINA: a novel energy scalable quality-aware color interpolation architecture. *Microprocess. Microsyst.* **67**, 8–17 (2019). <https://doi.org/10.1016/j.micpro.2019.02.011>
15. Given Imaging YouTube channel: <https://www.youtube.com/user/GivenImagingIntl/videos>. Accessed 28 July 2022
16. I.M. Gralnek, S.N. Adler, K. Yassin, B. Koslowsky, Y. Metzger, R. Eliakim, Detecting esophageal disease with second-generation capsule endoscopy: initial evaluation of the PillCam ESO 2. *Endoscopy* **40**(4), 275–279 (2008). <https://doi.org/10.1055/s-2007-995645>
17. Y. Gu, X. Xie, G. Li, T. Sun, D. Wang, Z. Yin, Z. Wang, Design of endoscopic capsule with multiple cameras. *IEEE Trans. Biomed. Circ. Syst.* **9**(4), 590–602 (2015). <https://doi.org/10.1109/TBCAS.2014.2359012>
18. Y. Gu, H. Jiang, X. Xie, G. Li, Z. Wang, in *An image compression algorithm for wireless endoscopy and its ASIC implementation*, IEEE Biomedical Circuits and Systems Conference (BioCAS), pp.103–106 (2016). <https://doi.org/10.1109/BioCAS.2016.7833735>
19. D.K. Iakovidis, R. Sarmiento, J.S. Silva, A. Histace, O. Romain, A. Koulaouzidis, X. Dray, in *Towards intelligent capsules for robust wireless endoscopic imaging of the gut*, In IEEE International Conference on Imaging Systems and Techniques (2014, October) pp. 95–100
20. G. Iddan, G. Meron, A. Glukhovsky, P. Swain, Wireless capsule endoscopy. *Nature* **405**, 417–418 (2000). <https://doi.org/10.1038/35013140>
21. M. Imran, K. Shahzad, N. Ahmad, M. O’Nils, N. Lawal, B. Oelmann, Energy-efficient SRAM FPGA-based wireless vision sensor node: SENTIOF-CAM. *IEEE Trans. Circuits Syst. Video Technol* **24**(12), 2132–2143 (2014). <https://doi.org/10.1109/TCSVT.2014.2330660>
22. J. Jang, H.J. Yoo, A capsule endoscope system for wide visualization field and location tracking. *IEEE Biomed. Circuits Syst. Conf. (BioCAS)* (2018). <https://doi.org/10.1109/BIOCAS.2018.8584812>
23. T.H. Khan, K. Wahid, Low-power and low complexity compressor for video capsule endoscopy. *IEEE Trans. Circuits Syst. Video Technol.* **21**(10), 1534–1546 (2011). <https://doi.org/10.1109/TCSVT.2011.2163985>
24. A. Koulaouzidis, D.K. Iakovidis, A. Karargyris, E. Rondonotti, Wireless endoscopy in 2020: Will it still be a capsule? *World J. Gastroenterol. WJG* **21**(17), 5119 (2015). <https://doi.org/10.3748/wjg.v21.i17.5119>

25. W.G. Kwack, Y.J. Lim, Current status and research into overcoming limitations of capsule endoscopy. *Clin. Endosc.* **49**(1), 8 (2016). <https://doi.org/10.5946/ce.2016.49.1.8>
26. G. Liu, G. Yan, S. Zhao, S. Kuang, A complexity-efficient and one-pass image compression algorithm for wireless capsule endoscopy. *Technol. Health Care* **23**(s2), 239–S247 (2015). <https://doi.org/10.3233/THC-150959>
27. G. Liu, G. Yan, B. Zhu, L. Lu, Design of a video capsule endoscopy system with low-power ASIC for monitoring gastrointestinal tract. *Med. Biol. Eng. Comput.* (2016). <https://doi.org/10.1007/s11517-016-1472-2>
28. N.V. Malathkar, S.K. Soni, Low complexity image compression algorithm based on hybrid DPCM for wireless capsule endoscopy. *Biomed. Signal Process. Control* **48**, 197–204 (2019). <https://doi.org/10.1016/j.bspc.2018.10.016>
29. N.V. Malathkar, S.K. Soni, A near lossless and low complexity image compression algorithm based on fixed threshold DPCM for capsule endoscopy. *Multimed. Tools Appl.* **79**(11), 8145–8160 (2020). <https://doi.org/10.1007/s11042-019-08347-w>
30. H.S. Malvar, L.W. He, R. Cutler, High-quality linear interpolation for demosaicing of bayer-patterned color images. *Proc. ICASSP* **3**, 485–488 (2004). <https://doi.org/10.1109/ICASSP.2004.1326587>
31. W. Marlicz, X. Ren, A. Robertson, K. Skonieczna-Żydecka, I. Łoniewski, P. Dario, G. Ciuti, Frontiers of robotic gastroscopy: a comprehensive review of robotic gastroscopes and technologies. *Cancers* **12**(10), 2775 (2020). <https://doi.org/10.3390/cancers12102775>
32. N. Merhav, G. Seroussi, M.J. Weinberger, Optimal prefix codes for sources with two-sided geometric distributions. *IEEE Trans. Inf. Theory* **46**, 121–135 (2000). <https://doi.org/10.1109/18.817513>
33. S.K. Mohammed, K.M. Rahman, K.A. Wahid, Lossless compression in Bayer color filter array for capsule endoscopy. *IEEE Access* **5**, 13823–13834 (2017). <https://doi.org/10.1109/ACCESS.2017.2726997>
34. R.F. Rice, *Some practical universal noiseless coding techniques*, Tech. Rep. JPL-79-22 (Jet Propulsion Laboratory, Pasadena, CA, 1979)
35. S. Schostek, M.O. Schurr, European research on wireless endoscopy—the VECTOR project, in *pHealth 2013*. (IOS Press, Amsterdam, 2013), pp.193–199
36. I. Shallari, M. O’Nils, From the sensor to the cloud: intelligence partitioning for smart camera applications. *Sensors* **19**(23), 5162 (2019). <https://doi.org/10.3390/s19235162>
37. P.R. Slawinski, K.L. Obstein, P. Valdastrì, Emerging issues and future developments in capsule endoscopy. *Tech. Gastrointest. Endosc.* **17**, 40–46 (2015). <https://doi.org/10.1016/j.tgie.2015.02.006>
38. P.H. Smedsrud, H. Gjestang, O.O. Nedrejord, E. Næss, V. Thambawita, S. Hicks, P. Halvorsen, (2020). Kvasir-Capsule, a video capsule endoscopy dataset. <https://doi.org/10.31219/osf.io/gr7bn>
39. R. Starosolski, New simple and efficient color space transformations for lossless image compression. *J. Vis. Commun. Image Represent.* **25**(5), 1056–1063 (2014). <https://doi.org/10.1016/j.jvcir.2014.03.003>
40. G.E. Tontini, F. Cavallaro, H. Neumann et al., Extensive small-bowel Crohn’s disease detected by the newly introduced 360 panoramic viewing capsule endoscopy system. *Endoscopy* **46**, 353–354 (2014). <https://doi.org/10.1055/s-0034-1377358>
41. P. Turcza, M. Duplaga, Hardware-efficient low-power image processing system for wireless capsule endoscopy. *IEEE J. Biomed. Health Inform.* **17**(6), 1046–1056 (2013). <https://doi.org/10.1109/JBHI.2013.2266101>
42. P. Turcza, M. Duplaga, Near-lossless energy-efficient image compression algorithm for wireless capsule endoscopy. *Biomed. Signal Process. Control* (2017). <https://doi.org/10.1016/j.bspc.2017.04.006>
43. P. Turcza, M. Duplaga, Energy-efficient image compression algorithm for high-frame rate multi-view wireless capsule endoscopy. *J. Real-Time Image Process.* **16**(5), 1425–1437 (2019). <https://doi.org/10.1007/s11554-016-0653-4>
44. Q. Wang, S. Chen, in *A low power prediction SAR ADC integrated with DPCM data compression feature for WCE application*, IEEE Biomedical Circuits and Systems Conference (BioCAS), pp.107–110 (2016). <https://doi.org/10.1109/BioCAS.2016.7833736>
45. M.J. Weinberger, G. Seroussi, G. Sapiro, LOCO-I, A low complexity context-based lossless image compression algorithm. *Proc. 1996 Data Comp. Conf.*, <https://doi.org/10.1109/DCC.1996.488319> (1996)
46. L.S. Xu, M.Q.H. Meng, C. Hu, Effects of dielectric values of human body on specific absorption rate following 430, 800, and 1200 MHz RF exposure to ingestible wireless device. *IEEE Trans. Inf. Technol. Biomed.* **14**(1), 52–59 (2010). <https://doi.org/10.1109/TITB.2009.2029853>

47. J. Xue, X. Xie, G. Li, Z. Wang, in *Guided Frequency Filter for Block-DCT Compressed Capsule Endoscopic Images*, IEEE Biomedical Circuits and Systems Conference (BioCAS), pp.1-4 (2018). <https://doi.org/10.1109/BIOCAS.2018.8584768>
48. D.C. Yates, A.S. Holmes, Preferred transmission frequency for size-constrained ultralow-power short-range CMOS oscillator transmitters, *IEEE Trans. Circuits Syst. I Regular Papers* **56**(6), 1173–1181 (2008). <https://doi.org/10.1109/TCSI.2008.2008506>
49. S. Yim et al., Biopsy using a magnetic capsule endoscope carrying, releasing, and retrieving untethered microgrippers. *IEEE Trans. Biomed. Eng.* **61**(2), 513–521 (2013). <https://doi.org/10.1109/TBME.2013.2283369>
50. G. Yingke, X. Xie, G. Li, T. Sun, D. Wang, Z. Yin, Z. Wang, in *Design of micro-ball endoscopy system*, IEEE Biomedical Circuits and Systems Conference (BioCAS), pp.208-211 (2012). <https://doi.org/10.1109/BioCAS.2012.6418460>
51. M.R. Yuce, T. Dissanayake, Easy-to-swallow wireless telemetry. *IEEE Microw. Mag.* **13**(6), 90–101 (2012). <https://doi.org/10.1109/MMM.2012.2205833>
52. Y. Zheng, L. Hawkins, J. Wolff, O. Goloubeva, E. Goldberg, Detection of lesions during capsule endoscopy: physician performance is disappointing. *Am. J. Gastroenterol.* **107**(4), 554–560 (2012). <https://doi.org/10.1038/ajg.2011.461>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.