Efficient Encoding of Wireless Capsule Endoscopy Images Using Direct Compression of Colour Filter Array Images

A Thesis Submitted to the College of
Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Electrical and Computer Engineering
University of Saskatchewan
Saskatoon, SK, Canada

by

Shahed Khan Mohammed

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan
Canada, S7N 5A9

Abstract

Since its invention in 2001, wireless capsule endoscopy (WCE) has played an important role in the endoscopic examination of the gastrointestinal tract. During this period, WCE has undergone tremendous advances in technology, making it the first-line modality for diseases from bleeding to cancer in the small-bowel. Current research efforts are focused on evolving WCE to include functionality such as drug delivery, biopsy, and active locomotion. For the integration of these functionalities into WCE, two critical prerequisites are the image quality enhancement and the power consumption reduction. An efficient image compression solution is required to retain the highest image quality while reducing the transmission power. The issue is more challenging due to the fact that image sensors in WCE capture images in Bayer Colour filter array (CFA) format. Therefore, standard compression engines provide inferior compression performance.

The focus of this thesis is to design an optimized image compression pipeline to encode the capsule endoscopic (CE) image efficiently in CFA format. To this end, this thesis proposes two image compression schemes.

First, a lossless image compression algorithm is proposed consisting of an optimum reversible colour transformation, a low complexity prediction model, a corner clipping mechanism and a single context adaptive Golomb-Rice entropy encoder. The derivation of colour transformation that provides the best performance for a given prediction model is considered as an optimization problem. The low complexity prediction model works in raster order fashion and requires no buffer memory. The application of colour transformation yields lower inter-colour correlation and allows the efficient independent encoding of the colour components.

The second compression scheme in this thesis is a lossy compression algorithm with a 4×4 integer discrete cosine transformation at its core. Using the statistics obtained from a large dataset of CE image, an optimum colour transformation is derived using the principal component analysis (PCA). The transformed coefficients are quantized using optimized quantization table, which was designed with a focus to discard medically irrelevant information. A fast demosaicking algorithm is developed to reconstruct the colour image from the lossy CFA image in the decoder. Extensive experiments and comparisons with state-of-the-art lossless image compression methods establish the superiority of the proposed compression methods as simple

and efficient image compression algorithm. The lossless algorithm can transmit the image in a lossless manner within the available bandwidth. On the other hand, performance evaluation of lossy compression algorithm indicates that it can deliver high quality images at low transmission power and low computation costs.

Acknowledgments

I would like to convey my heartiest gratitude to my supervisor, Professor Khan Wahid for supervising my work. Starting with a little background in image compression, I was able to acquire content knowledge and contribute to the advancement of the state-of-the-art research with his valuable guidance and continual encouragement. I am privileged to have the opportunity to work under his supervision, which immensely enriched my graduate experience.

I would like to thank Professor Rajesh Karki, Professor Mark Eramian, Professor Francis Bui and Professor Khan Wahid for offering me excellent courses at the University of Saskatchewan.

I would also like to thanks all my friends and colleagues in the Multimedia Processing and Prototyping laboratory at the University of Saskatchewan. They are Dr. Tareq Khan, Mehedi Hasan, Ravi Shrestha, Xuechao Zhang, Ali Melli, Farah Deeba, and KMM Rahman.

Finally, I would like to express my deepest gratitude and love to my parents, my parents-in-law, and my wife Farah for their unconditional love, care, and support at each step of my life.

Table of Contents

Page

List of Tables

Table Number Page Number

List of Figures

Figure Number Page Number

List of Abbreviations

ACD	Adaptive colour difference
AGR	Adaptive Golomb-Rice
AIQ	Algebraic integer quantization
BI	Bilinear interpolation
BPP	Bits per pixel
CCD	Charge couple device
CE	Capsule endoscopy
CEIC	Computationally efficient image compressor
CELC	Cost-efficient lossless compression
CFA	Colour filter array
CMOS	Complementary metal oxide semiconductor
CPSNR	Colour peak signal to noise ratio
CR	Compression rate
CT	Colour transformation
DCT	Discrete cosine transform
DPCM	Delta pulse coded modulation
DST	Discrete sine transform
DWHT	Discrete Walsh-Hadamard transform
DWT	Discrete wavelet transform
EPE	Energy packing efficiency
FPCD	Fully pipelined colour demosaicking
FPS	Frame per second
HDL	Hardware description language

HP	Hierarchical prediction
HQLI	High- quality linear interpolation
IDA	Iron deficiency anamia
KLT	Karhunen Loeve transform
LCG	Local coding gain
LICS	Lossless image compression system
LLIC	Lossless low-power image compressor
MEP	Median Edge prediction
OGIB	Obscure gastrointestinal bleeding
ORCT	Optimum reversible colour transformation
PCA	Principal component analysis
RCT	Reversible colour transformation
RF	Radio-frequency
RGB	Red-green-blue
RI	Residual interpolation
SSIM	Structural similarity
WCE	Wireless capsule endoscopy

Chapter 1 - Introduction

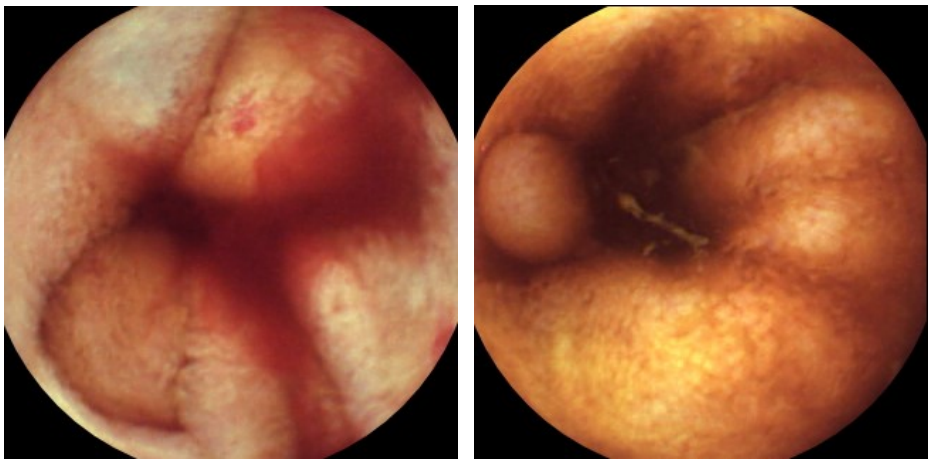
In 2001, “Given Diagnostic Imaging” launched its first commercial wireless capsule endoscopy system (WCE) i.e. PillCam. It was designed by an Israeli engineer “Gavriel Iddan” and British gastroenterologist “Paul Swain” [1][2]. WCE was a revolutionary invention because it allowed the non-invasive visualization of the small-bowel. Due to its capability of direct small-bowel visualization, WCE quickly became the first-line modality in the diagnosis of gastrointestinal diseases such as obscure gastrointestinal bleeding, unexplained iron deficiency, anaemia, Crohn’s disease and small-bowel polyps. Figure 1-1 shows examples of abnormalities captured by capsule endoscopy.

WCE offers patient-friendly alternative to wired endoscopy that caused discomfort and also required sedation in many cases. WCE brought a positive impact on the patient compliance thereby increasing the early detection of GI diseases such as cancer. As a result, WCE gained widespread popularity and several other companies now are marketing small-bowel endoscopic capsules such as PillCam SB3 [3], Endocapsule [4], OMOM Capsule [5], MiroCam [6] and Capsocam SV1 [7]. Table 1-1 summarizes the features of these systems [8]. In just a few years, WCE has become an invaluable tool in gastrointestinal disease management.

Despite continued technological advancements, WCE has several limitations such as poor image quality, passive locomotion and time-consuming manual reviewing.

Table 1-1: Comparison of video capsule endoscopes

	PillCam SB3	Endocapsule	OMOM Capsule	MiroCam	Capsocam SV1
Manufacturer	Given Imaging	Olympus	Jianshan	IntroMedic	CapsoVision
Length × Diameter	26mm × 11mm	26mm × 11mm	28mm × 13mm	25mm × 11mm	31mm × 11mm
Frame Rate (per second per camera)	2-6	2	3	3	3-5
Field of View (^{o})	156	160	140	170	360
Battery Life (h)	12	12	8	12	15



(a) (b)

Figure 1-1. Abnormalities captured in capsule endoscopy image (a) bleeding, (b) polypoid (courtesy: KID Dataset [9])

The main function of endoscopy is to inspect GI tract using imaging technique. A wired endoscope utilizes high-resolution image camera and light delivery systems powered by an external source. An optical fibre transmission system is used for transmitting power and receiving image data. On the other hand, the battery-operated WCE transmits the image data using radio-frequency (RF) transmission. As shown in Figure 1-2, a typical imaging system in WCE consists of five components i.e. a CMOS image sensor, an illumination system, an image processor, a RF transmitter and a power source [1]. The CMOS image sensor along with the illumination system constitutes the image acquisition system. The main function of the image processor is to compress the image data before transmission. Additionally, the image processor may contain pre-processing algorithms such as demosaicking and white balancing.

Using these modifications, WCE greatly alleviates the discomfort and pain of the patients caused by the cable in wired endoscopes. However, the hardware limitations brought up by these modifications such as power consumption and wireless transmission bandwidth, result in image-quality degradation [10]. As the battery occupies most of the space in capsule, frame rate and image resolution have been curtailed in the WCE to reduce the RF transmission power. Moreover, the available bandwidth for transmitting data through the human body is restricted to 2~3 Mbps by the Federal Communications Commission [11]. As shown in Table 1-1,

commercial WCEs typically yield images with QVGA (256x256) resolution at a maximum frame rate of 6 FPS. The low resolution of image often results in the unintentional omission of abnormal findings.

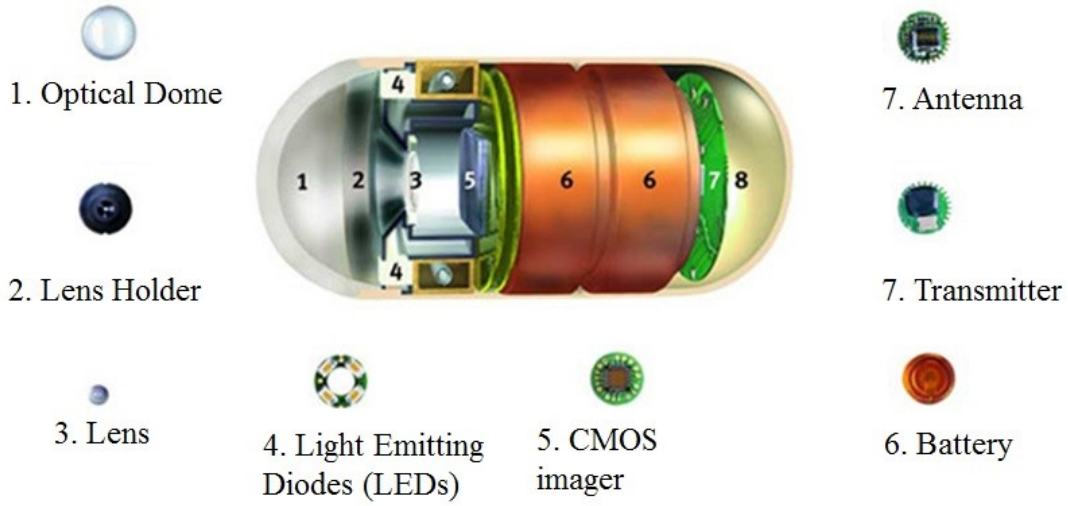


Figure 1-2. Visual demonstration WCE components (Courtesy: Given Imaging).

Besides diagnosis, image quality plays a vital role as a feedback signal for the remote manipulation system. A remote manipulation system is essential for allowing targeted biopsies or drug delivery. Most of the current commercial WCEs employ passive locomotion, where the natural peristalsis motion controls the capsule movement. As a result, WCE lags behind its wired counterpart in therapeutic procedures. In addition to increasing diagnostic yield, an efficient remote manipulation system can reduce the transit time, energy consumption and completion rate of the WCE [2]. It is foreseeable that the next generation capsule endoscopy system will replace the passive locomotion with active locomotion. The efficiency of remote manipulation system relies largely on the availability of image transmission with sufficient frame rate and high resolution [12].

One of the most challenging issues for the development of WCE is transmitting high-quality images while reducing the transmission power consumption. An efficient image compression system can effectively limit the transmission power by decreasing the amount of RF

transmission data. On the other hand, standard compression engines such as JPEG [13], JPEG-LS [14], JPEG-2000 [15] and JPEG-XR [16] are not suitable for WCE due to their high computational complexity. Moreover, these engines are not optimized for capsule endoscopy images. Therefore they cannot exploit the unique characteristics of capsule endoscopy images.

The purpose of this thesis is to propose and evaluate low complexity image compression system optimized for wireless capsule endoscopy system. The goal is to optimize the compression system to exploit the unique characteristics of capsule endoscopy images and provide the best compression performance while maintaining a low computational complexity and memory requirement. The development of a dedicated image-compression system for WCE will enable the next generation capsule endoscopy system to improve the image quality and incorporate functionality such as active locomotion and automated computer-aided detection system.

1.1 Previous Work on WCE Image Compression

In WCE, compression of CE image prior to transmission can reduce the power consumption by reducing the transmitted bits per pixel. This section summarizes the prior works on the compression of CE images. Due to space constraint in the capsule, the camera in WCE only has a single image sensor plane, consisting of either charge couple device (CCD) or a complementary metal oxide semiconductor (CMOS). WCE captures a colour image with this sensor by placing a colour filter array (CFA) in front of the image sensor. The result is a mosaic image where each pixel location contains either a red, blue or green samples. The full-colour image is reconstructed by applying a colour interpolation process called demosaicking, which estimates the two missing colour values at each pixel. The most straightforward approach to compression is to apply demosaicking prior to compression in order to generate full-colour image and design[17]–[21]. The major reason for the popularity of such demosaicking-first scheme is the standard compression engine such as JPEG, JPEG-2000, JPEG-XR and JPEG-LS, that are intended to compress still continuous tone colour image. However, such demosaicking-first approach (Figure 1-3 (a)) is found to be sub-optimum in the case of WCE due to several reasons. First, the demosaicking process adds irreversible distortion to the image. In addition, the demosaicking process requires computational resource and memory to interpolate the data and

the data to be compressed is tripled which increases the computational complexity and memory requirement of the compressor. To address these issues, various compression algorithms for direct coding of CFA image (Figure 1-3 (b)), that avoids the demosaicking stage prior to compression have been proposed for WCE.

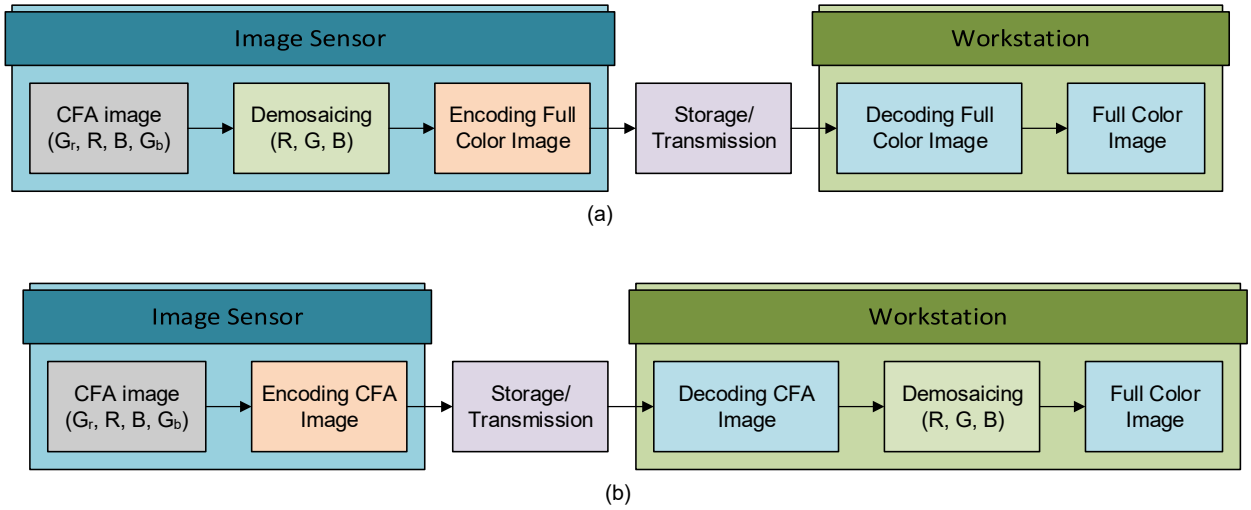


Figure 1-3. Block diagram of CFA image compression using: (a) demosaicking-first scheme; (b) compression-first scheme. In (a) the demosaicking occurs in the image sensor and encoder works on full-colour images. In (b) the demosaicking occurs in the decoder after the compression and decompression of the CFA image.

The state-of-the-art image compression algorithms for direct compression of CFA image can be broadly divided into two ways i.e. lossless [22] and lossy [12], [23]–[28]. In lossless compression, the image can be reconstructed without any loss of information in the decoder. Lossless compression plays a major role in reducing the error for both manual diagnosis and computer-aided decision system. On the other hand, lossy compression allows the addition of some distortion with a substantial increase in compression factor. The bottleneck in power consumption and transmission bandwidth force WCE to apply lossy compression. The major design concern in lossy compression is to preserve the diagnostic findings in the reconstructed image. Again these methods can be subdivided into prediction based and transform based coding.

1.1.1 Prediction Based Coding

In prediction based coding, a prediction model is used for predicting the current pixel by using the neighbouring pixel values. The prediction model employs a causal template to enable the decoder to reconstruct the pixel value. Among various prediction based compression schemes, JPEG-LS has been widely utilized for WCE [20], [25], [29]–[32]. Compared to other standard compression engines such as JPEG-2000 and JPEG, JPEG-LS is simple and has lower storage and computational requirement. These features make JPEG-LS attractive for real-time implementation in WCE. The block diagram of JPEG-LS is given in Figure 1-4. The prediction is computed based on the four neighbouring pixels as shown in Figure 1-4 (a, b, c and d). As the neighbouring pixels in CFA are sampled from different colour planes, they exhibit varying levels of pixel intensities. Therefore, direct compression of CFA image using JPEG-LS results in an inferior compression performance. Various proposed modification to address this issue are discussed below.

There are three popular modifications applied in the prior compression algorithm: i) structure separation, ii) structure rearrangement and iii) redefinition of prediction template.

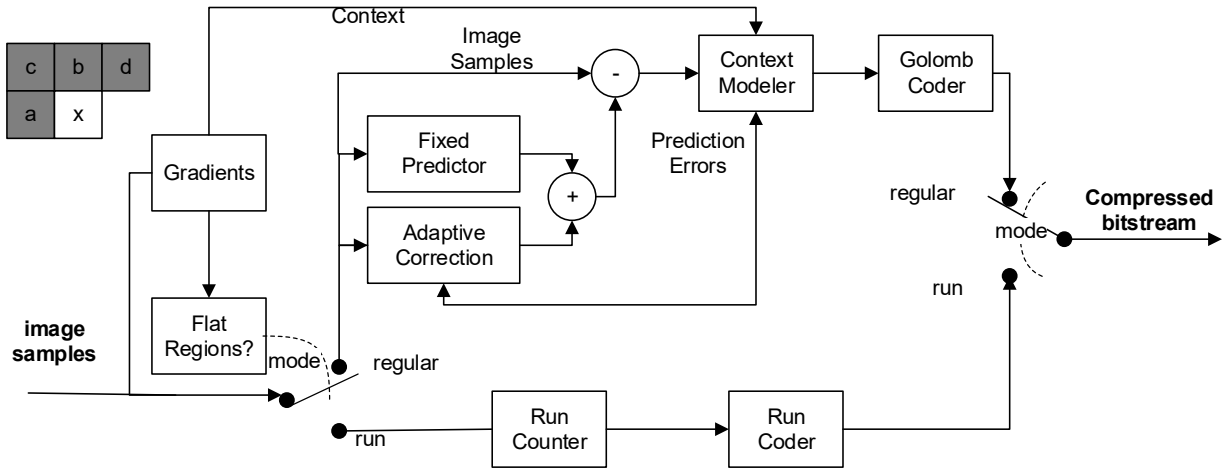


Figure 1-4. JPEG-LS Block Diagram based on LOCO-I algorithm [33]

Structure separation deinterleaves the different colour components into sub-images and then encodes them independently using the JPEG-LS engine. In WCE, the Bayer CFA data is commonly subdivided into green sub-images and non-green (red and blue) sub-images [34] [35]. The procedure is shown in Figure 1-5(a). This reduces the number of parallel paths thereby

leading to lower computational complexity. In order to reduce the artificial discontunities between R and B components, simple low-pass filtering is applied prior to compression. This low-pass filtering stage controls the trade-off between the compression rate and the distortion. As a result, this method cannot be applied in lossless compression as low-pass filtering stage is not a lossless process.

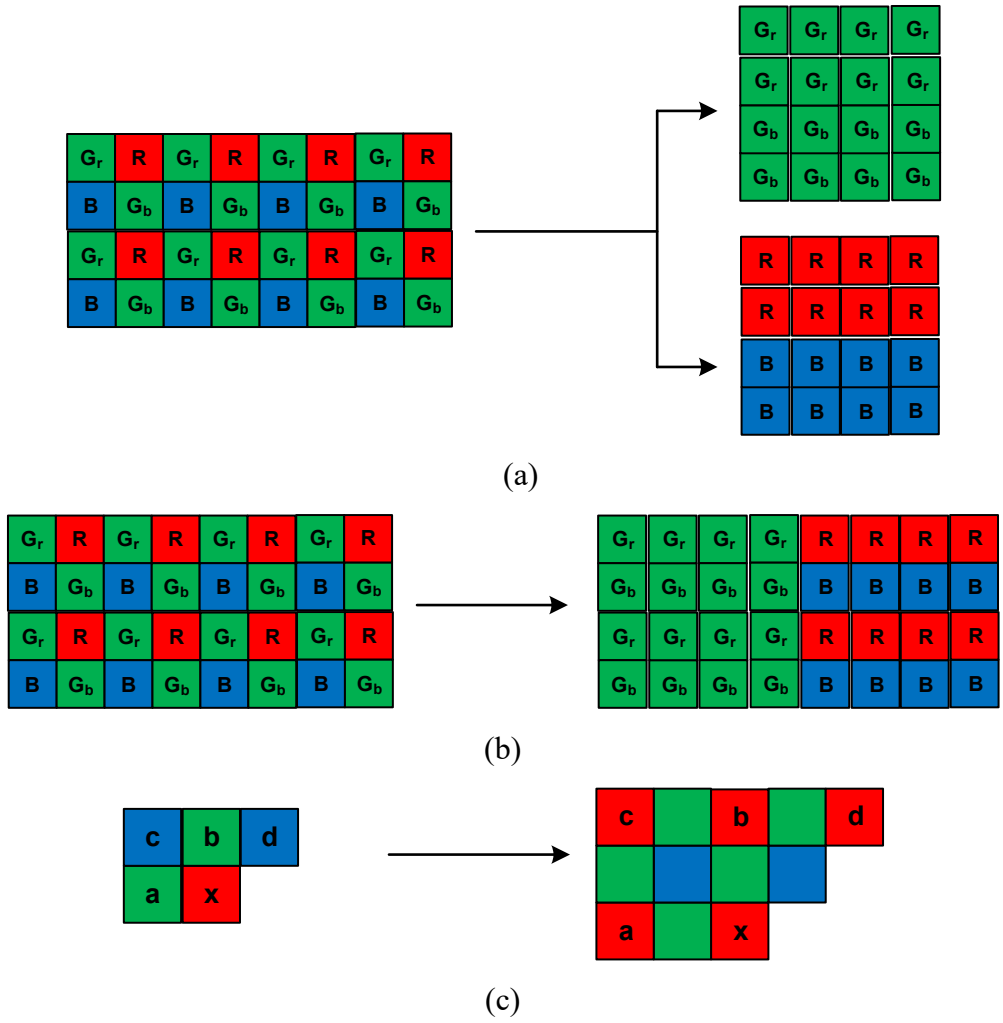


Figure 1-5. Different modification used in JPEG-LS based compression model for WCE are as follows: (a) Structure separation divides the CFA image into a green sub-image and a non-green sub-image, (b) Structure rearrangement shifts the green components to the left of the row and shifts the non-green components to the right and (c) redefining the prediction template to generate prediction from same colour components.

Instead of dividing the image into several sub-images, structure rearrangement shifts the elements into a suitable arrangement for JPEG-LS compression. For example: in [25], the green components are shifted to left and the non-green components to the right of each row (Figure 1-5(b)). The resulting mosaic image is then processed by a low-pass filter and fed directly to the JPEG-LS engine. This modification reduces the computational complexity compared to structure rearrangement [25]. However, both structure rearrangement and structure separation require frame memory for restructuring the components. The redefinition of prediction template for CFA image has been proposed in [32]. As shown in Figure 1-5(c), the work defines the position of a, b, c and d in the prediction template in order to generate the prediction using same colour pixels. The work also modified the JPEG-LS prediction model and proposed a hybrid entropy encoder optimized for CE image compression. As a result, the work has significantly reduced the computational complexity compared to previous works. This is the only previously published work on lossless compression of CE image.

1.1.2 Transform Based Methods

Transform coding is an efficient decorrelation technique for block-based image data, where the energy of the image data is condensed into a relatively small number of transformed coefficients. Due to their energy-packing efficiency, transform coding based on the discrete cosine transform (DCT) is widely used and is the core element in JPEG image compression. Most of the current low-bitrate lossy compression algorithm for WCE image follows a similar pipeline of JPEG. The JPEG of a colour image follows five basic stages as shown in Figure 1-6. First, a colour transformation reduces the spectral redundancy between the colour components. Then the individual colour components are divided into blocks that are transformed using a two-dimensional DCT. The transformed coefficients are quantized and then encoded using a coefficient encoder.

As direct compression of CFA image by JPEG produces periodic distortions [36], several modifications of JPEG algorithm have been proposed in the literature. In this section, a brief overview of various transform based coding proposed for WCE is given.

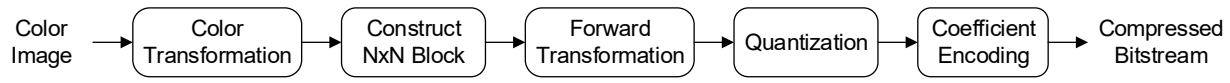


Figure 1-6. Block Diagram of Transform Coding

In order to reduce the artificial high frequencies in the CFA image due to the interlacing of different colour components, all the transform-based methods apply structure separation to divide the image into four sub-images. Each of the sub-images consists of pixels of a single colour plane. Most conventional colour conversion algorithms are designed for full-colour image and hence require three colour components. However, the CFA image contains two green components i.e. one red component and one blue component. Therefore, application of conventional colour conversion such as YCbCr would require demosaicking to generate the missing colour components, which will in turn increases the computational cost. Some compression algorithms [26]–[28] and [36] remove the colour transformation stage to avoid the application of demosaicking. In these cases, the image is first divided into four sub-images that consist of two G sub-images, one R sub-image and one B sub-image, each of which is one quarter the size of the original image. These sub-images are then transformed using a DCT transformation. The reduction of computational complexity of 2D DCT calculation requires approximation of DCT such as algebraic integer quantization based DCT [26] or integer transform [28]. The resultant transformed coefficients are quantized independently using optimized quantization table. Although these methods can yield a good compression ratio, the resulting peak signal to noise ratio is comparatively low (around 35dB).

In order to reduce the spectral redundancy between the sub-images in CFA, several approaches have been proposed [12], [37] and [38]. In [23], one of G sub-image is discarded and the other three sub-images are processed using a low complexity colour transformation, which is derived from YCbCr. The systems in [12] and [38] detail the reversible transformation introduced by Lee and Ortega [40] to map the CFA components into a new colour space. The conversion is applied on each 2x2 macro block of CFA image. It produces two luminance signals Y and two chrominance components Cb and Cr. The two Y components are merged to form a single Y sub-image with half the width of the original CFA image. The application of colour transformation significantly improves the compression performance resulting in compression

ratio of around 0.8 bits per pixel (BPP). In [12], the compression ratio is further increased by using the YYCoCg colour transformation and replacing the Huffman entropy encoder with a Golomb-Rice encoder.

1.2 Key Design Challenges

In the design of image compression, one of the key challenges is to find the optimum trade-off between performance and complexity. In addition, WCE operates under a set of unique resource constraints, including limitations on battery life, chip area, computational resources, memory and transmission bandwidth. This section summarizes the associated limitations and design constraints with image compression for wireless capsule endoscopy system.

- **Direct Compression of Raw Colour-Filter-Array Images:** In consumer camera, demosaicking is applied prior to the compression stage. However, the colour interpolation stage triples the data to be compressed resulting in a sub-optimum compression performance. Additionally, the sub-optimum demosaicking algorithm used by the in-chip processor introduces a significant amount of distortion in the data [41].

Therefore, WCE shifts the demosaicking algorithm to the receiver by directly compressing and transmitting the mosaic image. This approach avoids the increase in redundancy before compression and allows the use of high-complexity demosaicking algorithms for reconstructing the full-colour image. However, in CFA image, the intermixing of different colour pixel generates artificial discontinuities [36]. As a result, direct compression of mosaic images using standard continuous-tone image-compression algorithms such as JPEG, JPEG-2000 and JPEG-XR results in an inferior compression performance. The proposed compression algorithm should address this issue to facilitate efficient compression of raw CFA image for capsule endoscopy.

- **Image Quality**: The proposed compression algorithm should be able to yield high-quality images. Ideally, a lossless compression should be applied for transmission of the

endoscopic image without any distortion. For example: if the capsule is transmitting a VGA image (640 × 480, 8 bits per pixel) at a frame rate of 2 frames per second (fps), the lossless compression should reduce the bits per pixel (BPP) to 3.26 to 4.89 based on the available bandwidth for transmission (2-3 Mbps).

However, in view of facilitating s remote manipulation system, lossy compression is preferable [42]. According to a recent study, for a real-time remote-manipulation system to operate properly while complying with the FCC transmission bandwidth, the image should be compressed by a factor ranging between 5 to 20 [12]. Since lossless compression algorithm cannot provide such a high compression factor, a lossy compression algorithm must be applied. Particular attention should be given to retaining the medically-relevant information in the reconstructed image. Thus some distortion can be introduced but reconstructed images should have a very high image quality both in terms of objective and subjective measures.

- **Implementation Cost**: The computational complexity of the proposed compression scheme should be very low in order to reduce power consumption of the capsule. As the memory consumes a major part of the area and power in the compressor, the memory footprint of the proposed compression scheme should be small. Moreover, the compressor should work on raster order data so that it can be readily integrated with commercially-available image-sensors.

1.3 Thesis Objective

The objective of this work is to develop efficient image compression algorithms for WCE that works directly on the raw CFA image. To achieve this goal, the specific objectives of this thesis are as follows:

1. To improve the existing methods of colour transformation in prediction based lossless CFA image-compression for exploiting the redundancy between different colour components in CFA image and decreasing the BPP.

2. To develop a lossless compression algorithm to be used in WCE system with a comparable area and power to the state-of-the-art WCE systems and bits per pixel of 3.26 to 4.31. This enables transmission of full resolution image while complying with the transmission bandwidth available for RF transmission through the human body (2-3 Mbps).

3. To develop a lossy-compression algorithm for WCE to provide compression rate between 5-20 while keeping the peak signal ratio greater than 40 dB. The memory and computational complexity should be low to keep the area and power consumption low for WCE.

1.4 Thesis Organization

This thesis is organized in a manuscript-based style. The proposed methods and experimental results obtained are included in the form of submitted manuscripts. The remainder of the thesis is organized as follows. Chapter 1: provides the motivation, review of previous works, key design challenges and the thesis objective. Proposed optimum reversible-colour transformation models for lossless coding are presented in Chapter 2. Based on the proposed derivation model, a lossless image compression system for wireless capsule endoscopy system is proposed in Chapter 3. Proposed lossy image compression scheme is presented in Chapter 4. Finally, conclusion and directions for further research are given in Chapter 5.

Chapter 2 - A Lossless and Reversible Colour Space Transformation for Bayer CFA Images

The chapter includes a manuscript entitled ‘A Lossless and Reversible Colour Space Transformation for Bayer CFA Images’ by Shahed K. Mohammed, and Khan A. Wahid. The manuscript has been submitted to *IEEE Transaction on Circuits and Systems for Video Technology on January 2017.* In this chapter, a novel optimization method to derive the colour transformation for lossless prediction based coding is proposed. The method is evaluated by deriving colour transformation for natural image and comparing the performance with other colour transformation model. Later this optimization method is applied for deriving the colour transformation for CE image (Chapter 3).

Abstract—This chapter presents two low-complexity integer- reversible colour transformation for colour filter array (CFA) images, which are derived using a novel optimization method for reducing the prediction error variance and inter-colour correlation. These colour spaces can lower lossless bitrate of low complexity prediction model without using high complexity interpolation and inter-colour prediction scheme. Therefore, they are suitable for hardware implementation. One of the proposed colour transformation fully eliminates buffer memory requirement as it works directly on the raster order data. The second colour transformation, working on the 2x2 macroblock of colour components, requires storing only a half row of pixel values in memory and has an excellent decorrelation capability. In experiments, the proposed ORCTs reduced the computational complexity by 85% and buffer memory by 90% compared to state-of-art decorrelation technique used for CFA. These colour transformations can increase the prediction gain by 0.45 dB while significantly reducing the bits per pixel for standard lossless compression engine.

Index Terms—Colour filter array (CFA), reversible colour transformation (RCT), Lossless compression.

2.1 Introduction

Most digital cameras employ a single image sensor plane along with a red-green-blue (RGB) colour filter array (CFA), such as Bayer CFA (Figure 2-1 (a)), in which every pixel captures only one of the three primary colour components. The resultant image is called a mosaic image [43]. However, conventional compression engines, such as JPEG, are designed to work on the full-colour image; as a result, a demosaicking process reconstructs the full-colour image before compression. From the data compression perspective, this demosaicking stage triples the data to be compressed. Such data expansion is particularly questionable for lossless compression in applications such as image archiving and medical imaging [41]. An alternative approach where demosaicking occurs after compression offers simpler low power optical sensor as the computationally expensive process such as demosaicking can be shifted to powerful computer [44]–[48]. Specifically for lossless compression, these alternative schemes outperform the conventional methods. Among different types of lossless CFA compression methods, the

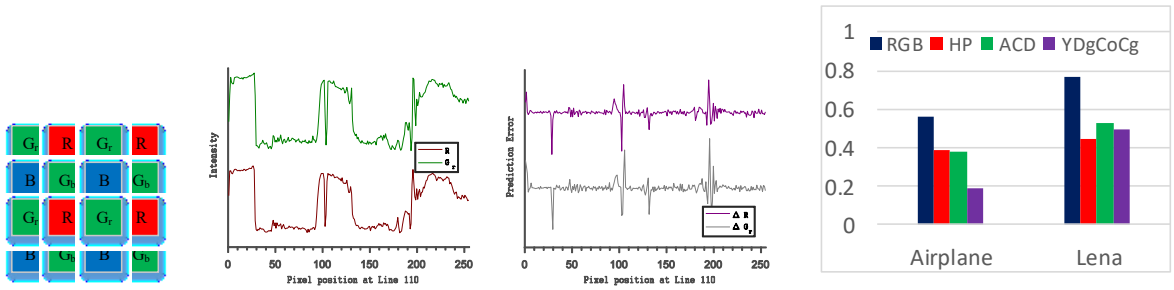
prediction-based approach using hierarchical prediction (HP) model [46] and adaptive colour difference (ACD) model [47] have demonstrated superior performance. These algorithms can provide a very promising lossless compression rate for CFA images by exploiting the inter-colour correlation. To this end, these models process the green (G) components first, then apply locally adaptive interpolation, and colour difference estimation in the non-green pixels. However, the associated requirement of a large buffer memory, computationally intensive interpolation, and inter-colour prediction process makes them unsuitable for hardware implementation in low-power image sensors. The alternative method of reducing the inter-colour correlation by applying a suitable colour transformation techniques have been widely utilized in lossy compression algorithm for mosaic images [36], [40]. But these methods cannot be incorporated in the lossless algorithms because associated colour transformations are not integer reversible, and hence cause rounding errors. In a recent work, a lifting-based reversible colour transformation YDgCoCg was presented for Bayer CFA image [48] which can substantially improve the lossless bitrate of transform-based methods, such as JPEG-2000 and JPEG-XR.

In this chapter, first, a novel derivation method is proposed to find the optimum colour transformation for a prediction based lossless CFA compression scheme (Section 2.2.1). This method incorporates the prediction model in the derivation of optimum colour transformation that can reduce the inter-colour correlation. Then, two optimum reversible colour transformation (ORCT) colour space for Bayer CFA image with varying implementation cost are proposed (Section 2.2.2). These colour spaces can efficiently improve the prediction based lossless performance by reducing the variance in error signal. Experiments with standard lossless engines such as JPEG-LS, JPEG-2000 and JPEG-XR demonstrated that the proposed ORCTs could be integrated with standard engines to improve the lossless bit rate with a negligible increase in computational complexities and buffer memories (Section 2.3).

2.2 Proposed method

The Bayer CFA pattern consists of repeating 2x2 macroblock of two G, one B, and one R components as shown in Figure 2-1(a). Throughout this thesis, the rows containing the R components and B components in the CFA image is denoted as R-line and B-line respectively;

the parameters associated with these lines are also distinguished by subscripts '*r*' and '*b*' respectively.



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

Figure 2-1. Spectral correlation in Bayer CFA image (a) Bayer CFA pattern for a 2x2 block, (b) Profile of the colour components G and R in the line 219 at Airplane CFA image which shows the presence of significant intra-colour and inter-colour redundancy (c) Profile of the residual signal in the same line obtained using a DPCM prediction model with reduced intra-colour redundancy but noticeable inter-colour redundancy, (d) Comparison of average inter-colour correlation in the prediction error signal for different colour transformations for airplane and Lena images

In CFA image, a significant inter-colour redundancy exists between the colour components. Error signals generated by individual processing of the colour components using a prediction model reduces the intra-colour redundancy. However, these error signal exhibits a significant inter-colour correlation. For example, the row profile R and G_r taken from one line in the Airplane image is plotted in Figure 2-1 (b). It is evident that a significant intra-colour redundancy or spatial redundancy and inter-colour redundancy exist between the components. A prediction model such as delta pulse coded modulation (DPCM) can effectively reduce the spatial redundancy as evident from the error signal ΔR and ΔG_r Figure 2-1 (c). However, there is still noticeable spectral redundancy between the residual signal from two different colour channels. Therefore, independent encoding of the colour component gives a sub-optimal compression performance as the spectral redundancy is not efficiently exploited. Application of colour transformation such as YDgCoCg or the colour difference signal estimation such as HP

and ACD reduces the spectral correlation in the original signal as well as in the residual error signal as evident from Figure 2-1 (d).

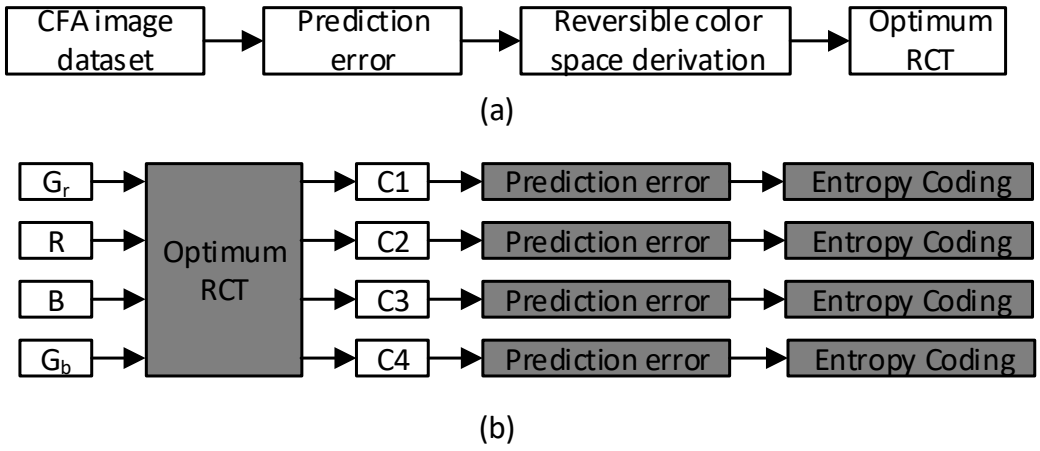


Figure 2-2. Overview of the proposed scheme (a) Optimum Colour space derivation, (b) Integration of the optimum RCT in prediction based lossless coding

On the other hand, the performance of prediction based image compression depends substantially on the variance of error signal [49]. It is expected that smooth images will have smaller error variance and can be coded with fewer bits. Therefore, the goal is to derive an optimization model that minimizes both the prediction error variance and the inter-colour correlation using a reversible colour transformation. First, the method is applied to obtain the colour transformation for both R-line and B-line, which can efficiently reduce the redundancy in the horizontal direction. Later, a separable colour transformation is obtained to lessen the redundancy in the vertical direction between R-line and B-line. An overview of the proposed method is given in Figure 2-2(a). The proposed colour transformation would improve the performance of prediction based CFA image coding systems that work on deinterleaved colour components separately (Figure 2-2 (b)).

2.2.1 Optimum Colour Transformation Derivation

It is seen from Figure 2-1 (a) that there are only two colour components in each line. The goal is to derive a lossless colour space defined by a 2x2-transformation matrix A which can

reduce the redundancy of the two colour components in raster order fashion. If the original colour components are P and Q, and the transformed colour components are W and D, the transformation can be written as:-

$$\begin{bmatrix} W \\ D \end{bmatrix} = A \begin{bmatrix} P \\ Q \end{bmatrix} \text{, where } A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{2-1}$$

To make the transformation reversible, there are two constraints:

1. A must be unimodular, i.e. determinant should be 1 [50]
2. Elments in one row of A must have integer coefficients [51].

The goal is to derive the optimum luminance and chrominance channel represented by a weighted average and difference of the inputs respectively. The sum of terms in the chrominance weight should be equal to zero, which ensures the gray colours have no chrominance value. Let the first row in A represents the weighted average while the second row represents the colour difference. This is reflected in the denotation of the transformed components in this thesis, where W represents the weighted average component and D refers to difference component. Considering the constraints for reversibility, the following two relationships can be formulated:

$$a_{11} + a_{12} = 1 \text{ and } \quad a_{21} + a_{22} = 0 \tag{2-2}$$

From the unimodular constraint and eqn (2-2), the condition to achieve lossless transformation can be found as:

$$|a_{21}| = |a_{22}| = 1 \tag{2-3}$$

So for each line, determining the ORCT translates to determining the coefficient a_{11} that will define the optimum colour space:-

$$A = \begin{bmatrix} a_{11} & 1 - a_{11} \\ 1 & -1 \end{bmatrix} \tag{2-4}$$

The goal is to minimize the lossless rate which largely depends on the variance of the error signal for prediction coding [49]. Therefore, the variance of the error signal is taken as the cost of the prediction model. Now if the error signal P, Q, W, and D are ΔP, ΔQ, ΔW, and ΔD respectively, then the relationship between the error signals can be established as:-

$$\Delta W = a_{11}\Delta P + (1-a_{11})\Delta Q \text{, and } \Delta D = \Delta P - \Delta Q \tag{2-5}$$

Then the prediction cost as the normalized error variance can be written as:-

$$J_{pred}(a_{11}) = \frac{1}{I^2_{\max}} \left(\begin{array}{l} \left(1+a_{11}{}^2\right)\frac{1}{N}\sum_{i=1}^{N}\Delta P^2(i) + \left(1+\left(1-a_{11}\right)^2\right)\frac{1}{N}\sum_{i=1}^{N}\Delta Q^2(i) + \\ 2\left(a_{11}(1-a_{11})-1\right)\frac{1}{N}\sum_{i=1}^{N}\Delta P(i)\Delta Q(i) \end{array} \right) \tag{2-6}$$

Here N denotes the number of samples and I_{max} denotes the maximum intensity value. To measure the inter-colour correlation, the Pearson correlation coefficient is calculated between the error signals from transformed colour space as below:-

$$J_{corr}(a_{11}) = Corr(\Delta W, \Delta D) = \frac{\sum_{i=1}^{N}\left(\Delta W(i)-\overline{\Delta W}\right)\left(\Delta D(i)-\overline{\Delta D}\right)}{\sqrt{\sum_{i=1}^{N}\left(\Delta W(i)-\overline{\Delta W}\right)^2 \sum_{i=1}^{N}\left(\Delta D(i)-\overline{\Delta D}\right)^2}} \tag{2-7}$$

Here N denotes the number of samples. These two cost functions are then combined using eqn (2-8)**:**

$$J^R_{Total} = J^R_{pred} + \lambda \times J^R_{corr} \tag{2-8}$$

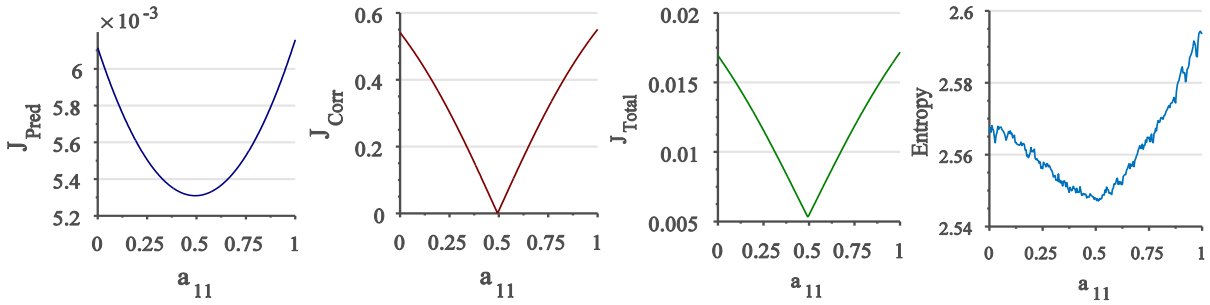


Figure 2-3. Variation of cost functions and entropy with respect to a_{11} for Hats image (a) Normalized Error Variance; (b) Correlation; (c) Total Cost (d) Entropy of Error Signal in all cases minima occurs at $a_{11} \approx 0.5$

Here, λ is the Lagrange multiplier that controls the trade-off between prediction error and correlation. Heuristically, the value of λ is set as 0.02 to keep the order of the two cost functions

compatible. To illustrate the cost function constitutes a convex optimization problem, the error signal of 'Hats' image in the Kodak dataset is calculated. Then using the error signal, J^R_{Total}, J^R_{corr}, J^R_{pred} as a function of a_{11} are plotted in Figure 2-3. The optimum value of a_{11} can be found from the local minima of Figure 2-3(c). The entropy of error signal is plotted in Figure 2-3(d) to reaffirm that minima of the cost function essentially provides the best compression performance since the minimum entropy index coincides with minima of J_{Total}. In the following section, two different ways are presented to derive colour space using the optimization method.

2.2.2 ORCT-1: Raster Order and Offline Optimization

In the first method, the optimum transformation matrix is determined separately for R-line and B-line according to the rules set out in eqns (6)-(8). The goal is to reduce the redundancy in the horizontal direction in raster order data. The optimization model is derived based on Kodak dataset in this experiment. Two prediction models are applied on the dataset: DPCM and Median Edge prediction (MEP). DPCM is a very efficient low-complexity prediction model used in multiple applications for lossless medical image coding [52]. On the other hand, MEP is the prediction model used in JPEG-LS, which can efficiently code both edge and smooth regions [14]. After generating the prediction error for R-line and B-line, an exhaustive search is applied to find the value of a_{11} that minimizes the cost function J_{Total}. To avoid division operation, the parameter is searched only on the integer multiple of $\frac{1}{256}$, which ensures a low complexity colour transformation matrix. The results for MEP and DPCM are shown in eqn (2-9) and (2-10) respectively:-

$$\begin{bmatrix} W_r \\ D_r \end{bmatrix} = \begin{bmatrix} \frac{123}{256} & \frac{133}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} Gr \\ R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} W_b \\ D_b \end{bmatrix} = \begin{bmatrix} \frac{122}{256} & \frac{134}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} G_b \\ B \end{bmatrix} \tag{2-9}$$

$$\begin{bmatrix} W_r \\ D_r \end{bmatrix} = \begin{bmatrix} \frac{125}{256} & \frac{131}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} Gr \\ R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} W_b \\ D_b \end{bmatrix} = \begin{bmatrix} \frac{120}{256} & \frac{136}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} G_b \\ B \end{bmatrix} \tag{2-10}$$

Since a_{11} is close to 0.5 in all cases, the coefficients can be further simplified, and so the parameters for both models become the same as shown below in eqn (2-11):

$$\begin{bmatrix} W_r \\ D_r \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} Gr \\ R \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} W_b \\ D_b \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} G_b \\ B \end{bmatrix} \tag{2-11}$$

To validate this approximation, the average entropy of the transformed signal is measured in Kodak dataset using eqns (2-9), (2-10) and (2-11), which are 6.6542 bpp, 6.6535 bpp, and 6.6548 bpp respectively, indicating an increase of only 0.02% due to the simplification. It shows that the simplification does not significantly reduce the coding performance and yet allows it to be implemented using a lifting structure (as shown in Figure 2-4). This enables simple and efficient hardware implementation with only adders and shifters.

Table 2-1: Comparison of entropy (bpp) for different methods

	Airplane	Baboon	House	Lena	Pepper	Avg
ACD[47]	5.40	7.33	6.08	6.62	7.36	6.56
HP [46]	5.74	7.58	6.33	6.82	7.46	6.79
ORCT-1	5.48	7.40	6.17	6.70	7.37	6.62

Table 2-2: Comparison of entropy (bpp) for Y and Green channels

	Airplane	Baboon	House	Lena	Pepper	Average
Green	6.79	7.47	6.53	7.59	7.49	7.18
W_r /W_b	6.64	7.42	6.39	7.37	7.44	7.05

In previous works on Bayer CFA [46] [47], similar difference components as given in equation (11) were used to make the signal smoother. However, instead of using the neighboring green pixel, these works utilized interpolated green value in the position of R and B to calculate the difference component. Table 2-1 compares the entropy of the difference signal using the interpolation (HP, ACD) and without using the interpolation (ORCT-1). From the table, it can be seen that the colour difference component estimated using interpolation has a reduction of entropy by 0.9% than the difference component in ORCT-1. As both HP and ACD transmit the green channel unaltered as the luminance channel, the entropy of green component, and the weighted average component W_r and W_b are compared in Table 2-2, which shows 2% reduction

in the entropy in the W channel in ORCT than the green component. Therefore, ORCT-1 can provide a better compression by reducing the bit rate for the weighted average channel. Furthermore, the interpolation step increases the computational complexity and buffer memory requirement (Section 2.3- Table 2-4).

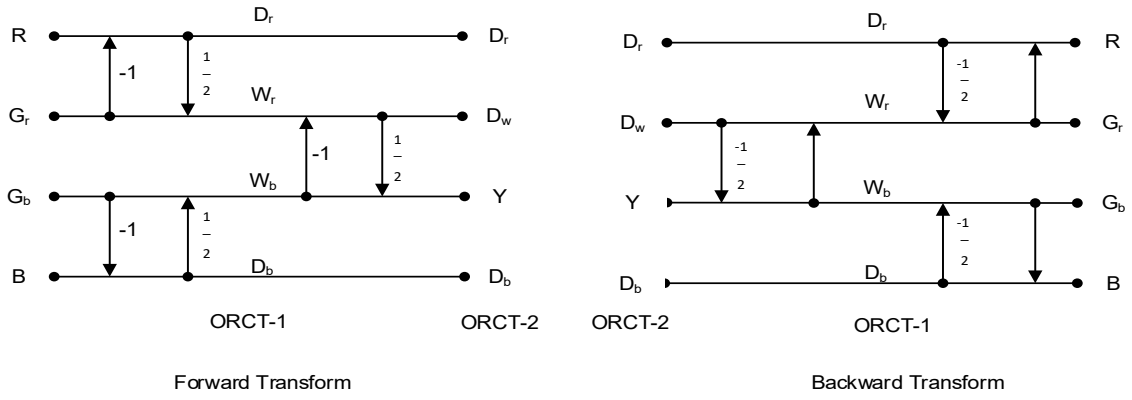


Figure 2-4. Flowgraph of the proposed reversible colour transformations ORCT-1 and ORCT-2.

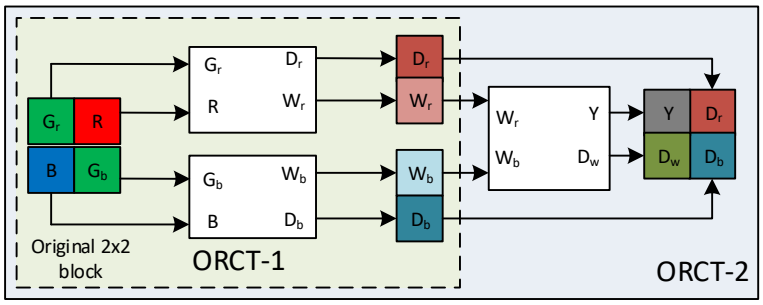


Figure 2-5. Block diagram of the proposed schemes

2.2.3 ORCT-2: Cascaded Offline Optimization

This method is an extension of ORCT-1. So far, R-line and B-line is considered as independent. However, the weighted average components of these two lines exhibit a strong correlation in the vertical direction(Fig.6-(g-i)). So a cascaded framework is proposed to exploit the redundancy between two W components in ORCT-1. The block diagram is shown in Figure 2-5.

In this case, W_r and W_b from ORCT-1 are taken as the input components for determining the error signal using MEP and DPCM. Then using the equation (6)-(8), the optimum colour space was determined. The results for MEP and DPCM are given respectively in eqn (2-12) and (2-13):-

$$\begin{bmatrix} Y \\ D_w \end{bmatrix} = \begin{bmatrix} \frac{122}{256} & \frac{134}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} W_b \\ W_r \end{bmatrix} \tag{2-12}$$

$$\begin{bmatrix} Y \\ D_w \end{bmatrix} = \begin{bmatrix} \frac{138}{256} & \frac{118}{256} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} W_b \\ W_r \end{bmatrix} \tag{2-13}$$

The transformation models can be further simplified by taking a_{11} approximated to 0.5. Then the original 2x2 block in Bayer pattern can be easily converted to a lossless integer transform by the procedure defined in equation (2-14). The corresponding signal flow graph is shown in Figure 2-4:-

$$\begin{aligned} \boldsymbol{D_r} &= R - G_r \\ W_r &= G_r + \left\lfloor \frac{D_r}{2} \right\rfloor \\ \boldsymbol{D_b} &= B - G_b \\ W_b &= G_b + \left\lfloor \frac{D_b}{2} \right\rfloor \\ \boldsymbol{D_w} &= W_r - W_b \\ \boldsymbol{Y} &= W_b + \left\lfloor \frac{D_w}{2} \right\rfloor \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} W_b &= Y - \left\lfloor \frac{D_w}{2} \right\rfloor \\ W_r &= W_b + D_w \\ G_b &= W_b - \left\lfloor \frac{D_b}{2} \right\rfloor \\ B &= D_b + G_b \\ G_r &= W_r - \left\lfloor \frac{D_r}{2} \right\rfloor \\ R &= D_r + G_r \end{aligned} \tag{2-14}$$

Forward Transform *Backward Transform*

Figure 2-6 shows examples of the application of ORCT-1 and ORCT-2 by expressing the colour components in quarter size sub-image. The third row is the sub-images generated by ORCT-1, where the redundancy is much lower than RGB sub-images. Finally, the fourth row shows the sub-images produced by ORCT-2, where redundancy between the weighted average components is minimized.

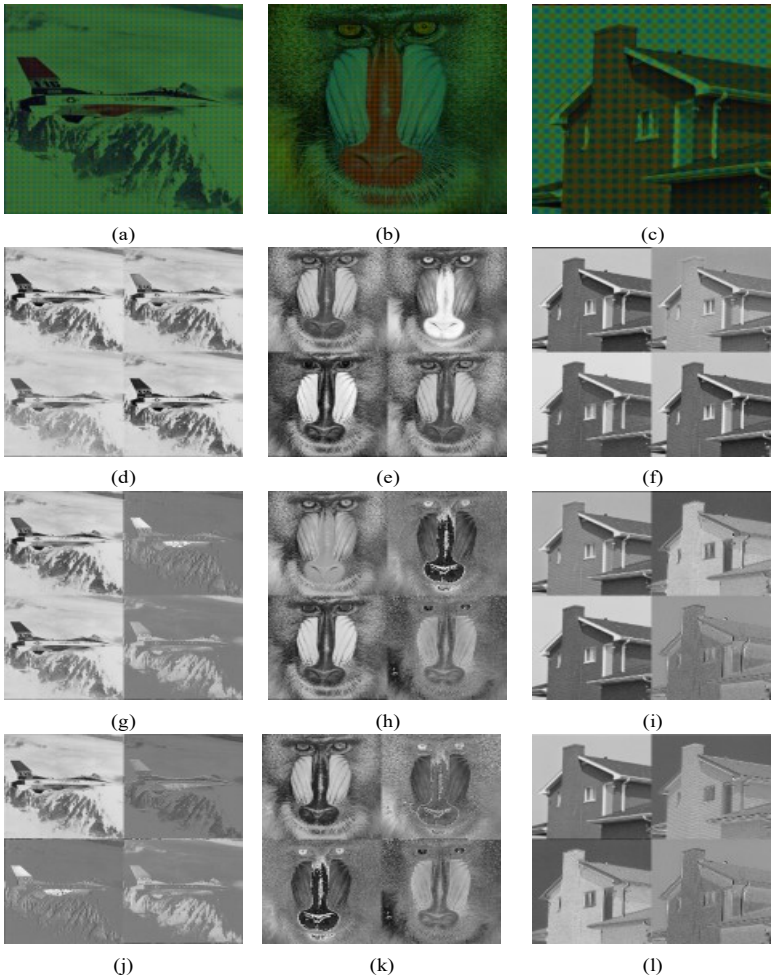


Figure 2-6. Visual Comparison of ORCT-1 and ORCT-2 : (a)-(c) Bayer CFA image, (d)-(f) Deinterleaved RGB components in clockwise order R-G_b-B-G_r , (g)-(i) Deinterleaved components in ORCT-1 in clockwise order C_r-C_b-Y_b-Y_r , (j)-(l) Deinterleaved components in ORCT-2, clockwise order C_g-C_b-C_r-Y

2.3 Experimental Results

To evaluate the performance of the proposed colour transformation, several experiments are conducted using the simulated CFA images from Kodak, Olympus E-P1, and NIKON D90 datasets[46]. For raw CFA images, the REAL-CFA database presented in [46] is used. The results are compared with original RGB colour space, YDgCoCg [48], HP [46] and ACD [47].

First, the performance of ORCT-1 and ORCT-2 is evaluated in reducing the entropy and average inter-colour correlation from the error signal. The results are shown in Table 2-3. Here the DPCM and MEP represents the error signal generated by these prediction models. As expected, the colour space ORCT-1 and ORCT-2, both reduce the inter-colour correlation and entropy in the error signal as well as in the original signal.

Table 2-3: The comparison between Correlation and Entropy in Kodak Dataset

		RGB	ORCT- 1	ORCT-2
Original	Correlation	0.81	**0.36**	**0.28**
	Entropy	7.11	***6.65***	***6.39***
DPCM	Correlation	0.48	***0.23***	***0.16***
	Entropy	5.78	***5.56***	***5.49***
MEP	Correlation	0.28	***0.11***	***0.09***
	Entropy	5.75	***5.52***	***5.49***

Second, the complexity of the different methods in terms of an average number of operations required per pixel is listed in Table 2-4. Operations that are considered are addition (A), multiplication, comparison, bit-shift, and measuring the absolute value (B). The memory is calculated as the minimum number of rows required to store to do the corresponding colour transformation in a raster order Bayer CFA data.

Table 2-4: Complexity and Memory requirement for different methods

Method	A	M	C	S	B	Total	Mem
YDgCoCg[48]	1.5	0	0	0.75	0	2.25	1
HP[46]	18	1.5	0	1	7	27.5	5
ACD[47]	6	2.5	1	0.5	5	15.0	5
ORCT-1	1	0	0	0.5	0	1.50	0
ORCT-2	1.5	0	0	0.75	0	2.25	1/2

A – adder; M – multiplier, C- comparator, S- shift, B- absolute value; Mem - memory

Both HP and ACD methods require a high number of operation including floating multiplication and division, and a significant amount of memory requirement. It is evident that ORCT-1 has the lowest complexity and memory requirement as it can work directly on the raster

order data. In ORCT-2, as the cascaded stage only requires the luminance value generated in the first stage, the line buffer is reduced to half than that of YDgCoCg.

Table 2-5: Prediction Gain for different colour transformations (two best results are marked in bold)

Model	Database	RGB	YDgCoCg [48]	HP [46]	ACD [47]	ORCT-1	ORCT-2
DPCM	Kodak Dataset	7.70	**7.92**	7.77	7.85	7.85	**8.15**
	Real CFA	12.91	**13.57**	13.14	13.13	13.18	**13.60**
	D90	15.65	**17.01**	15.57	15.92	16.49	**17.08**
	E-P1	13.83	**15.39**	14.46	12.57	14.64	**15.37**
	Standard Image	7.45	**7.94**	6.63	7.07	7.53	**7.91**
MEP	Kodak Dataset	8.08	8.06	7.94	8.57	**8.58**	**8.77**
	Real CFA	13.56	**14.22**	13.50	14.21	14.06	**14.48**
	D90	16.93	**18.13**	16.74	16.90	18.11	**18.70**
	E-P1	14.95	**16.53**	15.59	13.19	16.16	**16.93**
	Standard Image	9.34	9.43	7.71	8.52	**9.64**	**9.91**

In the next experiment, the performance of different methods in reducing the prediction error is measured using prediction gain (G_p). Prediction gain is defined as the ratio between the variance of the original signal, and the variance of the error signal and is given in eqn (2-15) [49]. The higher the prediction gain, the lower the variance of the error signal.

$$G_p = 10\log\frac{\sigma_{ori}}{\sigma_{pred}} \tag{2-15}$$

Here, σ_{ori} and σ_{pred} represent the variance of the original signal and the variance of the prediction error in transformed colour space respectively. It is seen from

Table 2-5 that using a colour transformation improves the prediction gain, signifying a reduction in the variance of the signal. It can be seen that YDgCoCg, ORCT-1, and ORCT-2 methods provide better results than applying the colour difference signal. The second method, ORCT-2, provides the best performance consistently in all cases since it reduces the variance by decorrealting the colour components in both vertical and horizontal directions.

Table 2-6: Average compression ratios (Bits per pixel) for different Datasets (Two best results are marked in bold)

Compression	Image Datebase	RGB	YDgCoCg [48]	HP [46]	ACD [47]	ORCT-1	ORCT-2
JPEG 2000	Kodak_Dataset	5.24	**5.11**	5.26	5.23	5.09	**5.03**
	Real CFA	4.54	**4.47**	4.60	4.58	4.53	**4.49**
	D90	3.97	**3.84**	3.97	4.03	3.88	**3.81**
	E-P1	4.35	**4.12**	4.40	4.33	**4.12**	**4.01**
	Standard Image	5.44	5.47	5.58	5.40	**5.36**	**5.33**
JPEG LS	Kodak_Dataset	5.03	**4.93**	5.06	5.05	**4.93**	**4.89**
	Real CFA	4.36	**4.34**	4.45	4.39	4.39	**4.32**
	D90	3.85	**3.70**	3.85	4.08	3.77	**3.72**
	E-P1	4.23	**3.97**	4.20	4.21	3.99	**3.89**
	Standard Image	5.24	5.26	5.32	5.30	**5.16**	**5.15**
JPEG-XR	Kodak Dataset	5.46	5.28	5.45	**5.21**	5.29	**5.16**
	Real CFA	4.81	**4.66**	4.82	4.68	4.75	**4.66**
	D90	4.28	4.06	4.23	**4.04**	4.13	**4.04**
	E-P1	4.68	4.40	4.70	**4.38**	4.43	**4.30**
	Standard Image	5.69	5.65	5.82	5.64	**5.59**	**5.54**

In the final experiment, the performance of the proposed colour transformation with standard compression codecs such as JPEG-2000, JPEG-LS, and JPEG-XR is compared. In each codec, the colour components are separated and compressed individually. The average compression rate in bpp is shown in Table 2-6. Like before, the proposed method ORCT-2 with cascaded colour transformation outperforms all other colour transformation methods.

2.4 Conclusion

In this chapter, two reversible colour transformation models based on prediction error of input signal are presented. These models take advantage of inter-channel correlations of four Bayer colours. Both colour transformations are lossless and cost much less hardware compared with many other methods. Experimental results show that both ORCT-1 and ORCT-2 colour spaces can effectively reduce the spectral correlations and improve the compression performance on the deinterleaved colour space.

Chapter 3 - YLMN Based Lossless Image Compressor for WCE

The chapter includes a manuscript entitled 'YLMN Based Lossless Image Compressor for Wireless Capsule Endoscopy' by Shahed K. Mohammed, KMM Rahman, and Khan A. Wahid. The previous chapter proposed an optimization method to derive reversible and lossless colour transformation for CFA image. In this chapter, the optimization method is applied to derive optimum colour transformation for CE image, which works as the core component of the proposed lossless compression algorithm.

Abstract— This chapter introduces a lossless colour filter array (CFA) image compression algorithm for wireless capsule endoscopy (WCE) system. The proposed pipeline consists of a novel YLMN colour transformation, a raster order prediction model, and a single context adaptive Golomb-Rice encoder to code the residual signal with variable length coding. An optimum reversible colour transformation derivation model is presented, which incorporate the prediction model to find the optimum colour transformation. This model has derived the YLMN colour transformation, which is optimized for capsule endoscopic image compression for the raster order prediction model. After the colour transformation, each colour component has been independently encoded with the low complexity raster order prediction model and Golomb-Rice encoder. The proposed algorithm is realized by VLSI technique, which shows the proposed algorithm can reduce the gate count by 38.9% and the memory requirement by 71.2% compare to the state-of-the-art. Extensive experimentation is performed using the full resolution sample capsule endoscopic image to validate the performance. The experiment illustrates the proposed algorithm can outperform existing lossless and near-lossless compression algorithm for capsule endoscope.

Keywords—Wireless Capsule Endoscopy, Colour Filter Array, Lossless Image Compression, Reversible Colour Transformation.

3.1 Introduction

Wireless capsule endoscopy (WCE) is a small pill size camera which is swallowed by the patient. The camera travels the gastrointestinal tract in a similar fashion to ingested food, captures images along the entire journey and sends these images wirelessly to the recorder outside [53]. WCE has been playing a pivotal role in gastrointestinal disease management by allowing non-invasive patient friendly monitoring of gastrointestinal tract. Particularly, WCE serves as the first line modality in the diagnosis of obscure gastrointestinal bleeding (OGIB), unexplained iron deficiency anaemia (IDA), small-bowel mucosal lesions, Chron's disease and Celiac disease [2]. With researches going on the development of next-generation wireless capsule with tools to perform biopsies, drug-delivery and active locomotion , WCE has the potential to replace diagnostic standard endoscopy within the next 15 years [54]. However, the

bottleneck of both commercial and research capsule is the image quality, image transmission power and the battery life [55]. Preserving the image quality by the use of lossless image compression could yield better performance from both manual and computer assisted reviewing process [42]. The self-powered battery and the long transit time limits the image resolution and the frame rate of the wireless capsule endoscopy. Therefore, keeping the image quality high while reducing the transmission power low and within the limitation of available bandwidth is a daunting challenge for the image compression system.

The image sensors in wireless capsule endoscopy mostly utilize Bayer colour filter array (CFA) to capture the colour information with only one image sensor plane [56]. A pre-processing method called demosaicking stage transfer the CFA image into a full-colour image through interpolation of the missing colour information. In compression perspective, demosaicking stage increases the redundancy in the image without adding any information. On the other hand, application of a suboptimal low complexity demosaicking stage severely degrades the image quality by introducing artificial colour artifacts particularly in the edge regions [41]. In wireless capsule endoscopy, this leads to degradation of critical findings in the mucosa layer, villous pattern, aphthous ulceration and underlying blood veins [57], [58]. A lossless compression of the raw CFA data allows the use a high complexity edge preserving demosaicking algorithm in the decoder after the transmission. On the other hand, as these images contain medical diagnostics, archiving them according to Picture Archiving and Communication system requires the lossless distortion [59]. However, standard lossless compression engine such as JPEG-LS, JPEG-2000, JPEG-XR or HEVC intra coding are not feasible due to their computationally expensive design and inferior performance in coding CFA image. Therefore a dedicated lossless compression for endoscopic images is sought in many of the works.

There were several lossless compression for wireless capsule endoscopy in the recent years [19], [59]. However, this method operates on the full-colour image and ignores the error generated and complexity added in the demosaicking stage. There are few near-lossless compression methods proposed for CFA endoscopic images that utilize JPEG-LS lossless algorithm along with a deinterlacing filter to independently code the different colour components in CFA image [20], [25], [34] . However, these algorithm requires a buffer memory for storing the context model. As well as the correlation between colour components in not sufficiently exploited in this algorithm. Recently, Malvar et. al. has proposed an optimum colour

transformation termed as YDgCoCg based on Karhunen Loeve Transform of natural CFA image [60]. By exploiting the inter-colour correlation in the CFA image through colour transformation, this algorithm has demonstrated an inspiring lossless compression rate for transform based code such JPEG-2000 and JPEG-XR. Similarly, an optimum colour transformation that can capture the unique characteristics in endoscopic CFA image can lead to a better image compression with a simple encoder.

In this chapter, a low complexity image compression system for CFA endoscopic image is proposed based on the optimization model proposed in Chapter 2. This derivation model is utilized to give a separable colour space transformation for completely exploiting the inter-colour correlation between the colour components in CFA image. Since the colour transformation is separable, it can work directly on the raster order data and leads to a very efficient circuit. Based on the lossless colour space transformation, a low complexity image compression system for wireless capsule endoscopy system is proposed. This compression system can significantly reduce the transmission power while transmitting the image in a lossless manner to the decoder. The proposed encoder uses a simple delta pulse coded modulation based prediction model with a low complexity adaptive Golomb rice encoder to entropy code the residual signal. The exclusion of high complexity prediction models such as median edge prediction, context modeling, and Huffman encoder lower the hardware implementation cost in terms of computational resource and memory.

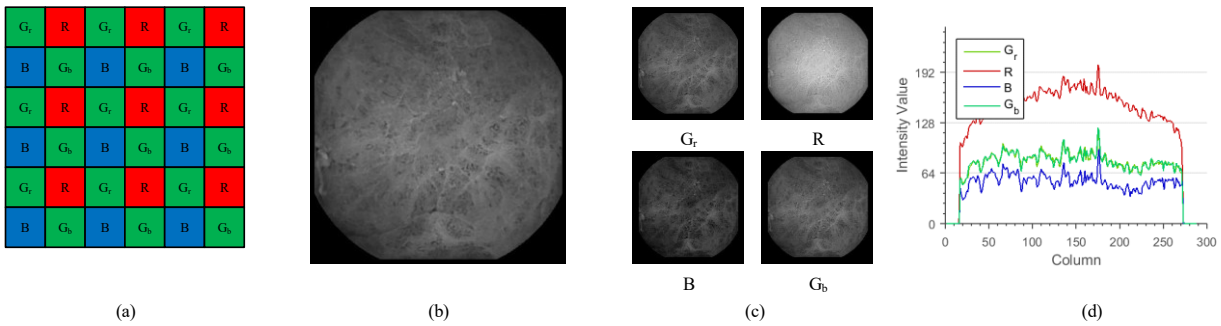


Figure 3-1. A typical CFA image in capsule endoscopy, (a) Bayer Pattern, (b) Mosaic Image, (c) Sub-images generated by separating the colour components, (d) Line plot showing the colour channel along row 100 in each sub-image to show the inter-colour redundancy.

In the experiments, the proposed method is compared with different lossless compression algorithms for both RGB and CFA endoscopic image, which demonstrates the best performance among the existing method and standard lossless engine. Performances are evaluated for some simulated CFA image of high-quality endoscopic images, as well as low contrast WBI images with various conditions such as ulcer, celiac disease. The compression shows that the proposed method outputs the lowest lossless bit rate on all the images.

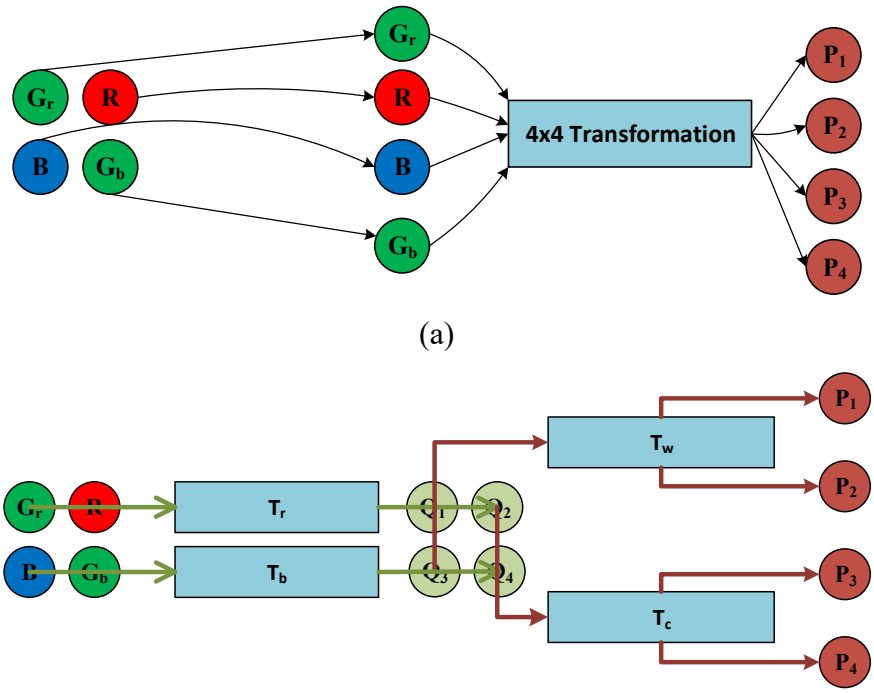


Figure 3-2. Workflow of different colour transformation: (a) Non-separable and (b) Separable

3.2 Mathematical Derivation of ORCT Derivation:-

A natural way to compress the CFA image generated by for example with Bayer pattern as shown in Figure 3-1(a) is to separate the image into several sub-images consisting of the different colour space. For example, by taking a pixel from each 2x2 macroblock, four sub-images comprising two green images, one red image, and one blue image can be generated. These sub-images exhibit a high degree of correlation between them similar to the correlation between the colour channel in RGB image. For example, row 100 for each sub-image has been

plotted in Figure 3-1 (d), showing that the similarity in the colour components. The goal of the colour space transformation is to reduce the redundancy between these sub-images and improve the performance of prediction based coding. To this end, the optimum colour derivation method presented in Chapter 2 is applied for deriving the colour space transformation for CE image.

3.2.1 ORCT for WCE

Deriving image-wise ORCT will introduce a substantial increase in computational overhead in the capsule compression system. Instead, an extensive endoscopic image dataset is employed to derive the optimum colour transformation in offline. The dataset contains 200 images taken for different condition and location in capsule endoscopy from Gastrolab [61]. To derive the colour transformation, the model is first applied for deriving the row transformation matrix for R-line and B-line in the CFA image. The residual error signal was determined using the DPCM predictor. The results are shown in eqn:-

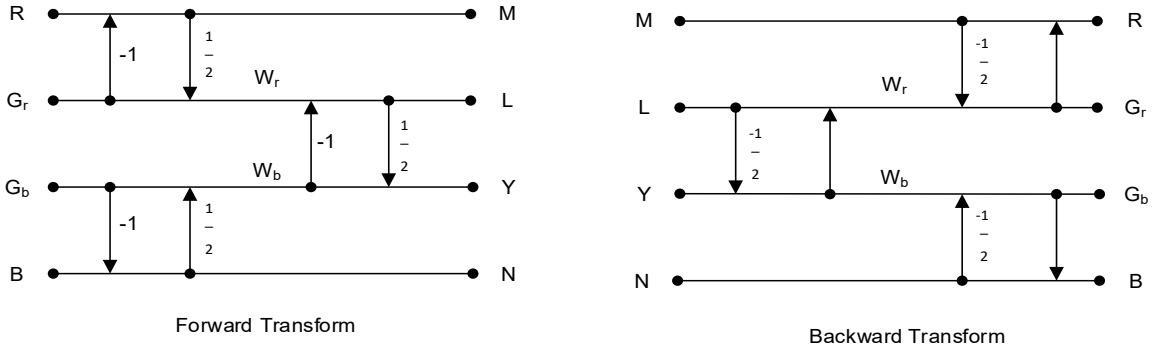


Figure 3-3. Flowgraph of the proposed reversible YLMN colour transformation

$$\begin{bmatrix} W_r \\ D_r \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & 1 \end{bmatrix} \begin{bmatrix} G_r \\ R \end{bmatrix} \tag{3-1}$$

$$\begin{bmatrix} W_b \\ D_b \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} G_b \\ B \end{bmatrix} \tag{3-2}$$

To calculate the column transformation matrix, the output of the row transformation matrix is fed to the derivation model as inputs. The results showed that for difference signal, the original signal contains the lowest entropy. This indicates that R-G and G-B have a lower correlation. However, the weighted signal contains a high amount of correlation, which can be reduced by using the column transformation matrix. So the column transformation is shown in eqn (3-3) and (3-4).

$$\begin{bmatrix} Y \\ L \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} W_r \\ W_b \end{bmatrix} \tag{3-3}$$

$$\begin{bmatrix} M \\ N \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} D_r \\ D_b \end{bmatrix} \tag{3-4}$$

The colour transformation is denoted as YLMN. The signal flow graph of the proposed lossless YLMN colour transformation is shown in Figure 3-3. Figure 3-4 shows the four resultant sub-images after the ORCT colour transformation for the image given in Figure 3-1.

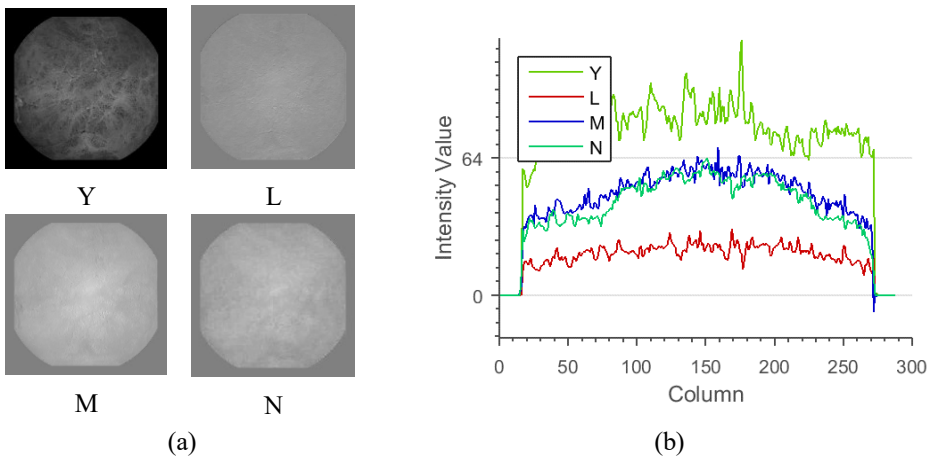


Figure 3-4. Colour decomposition in the proposed YLMN colourspace. (a) The colour components, (b) The line plots showing the row 100 in each sub-images to demonstrate the reduction in correlation between the transformed colour space

3.3 YLMN Based Lossless Image Compression for WCE

Figure 3-6 illustrates the block diagram of the proposed lossless encoder based on the proposed YLMN colour transformation. The encoder consists of YLMN colour transformation, structure separation, a simple delta pulse coded modulation prediction model and a single context adaptive Golomb-Rice encoder. For the input Bayer CFA image shown in Figure 3-1(b), first each R-line and B-line is transformed using the row transformation matrices defined in eqn (3-1) and (3-2) respectively. The difference signal generated in this step are sent directly to the prediction model. On the other hand, the weighted average signals are stored in buffer memory for further processing using the column transformation matrix defined in eqn (3-3). This transformation reduces the inter-colour redundancy between the colour components. After the colour transformation, a structure separation stage divides the images into four sub-images, where each sub-image contains pixels from one colour-plane. This step removes the artificial discontinuities between the pixels. Then each sub-images pass through a DPCM prediction model, which produces the prediction error signal dY, dL, dM and dN. Note that, the prediction model works in raster order fashion and does not require any line buffer. Therefore dM and dN, which are generated in raster order fashion, can be processed in raster order fashion. The details of the colour transformation, structure separation, and prediction model will be discussed in this section. Finally, a low complexity but efficient Golomb-Rice Encoder code the error signal generated by the prediction model to produce the bit stream. The bit stream is then sent wirelessly using the RF transmitter to the data recorder located outside of the body.

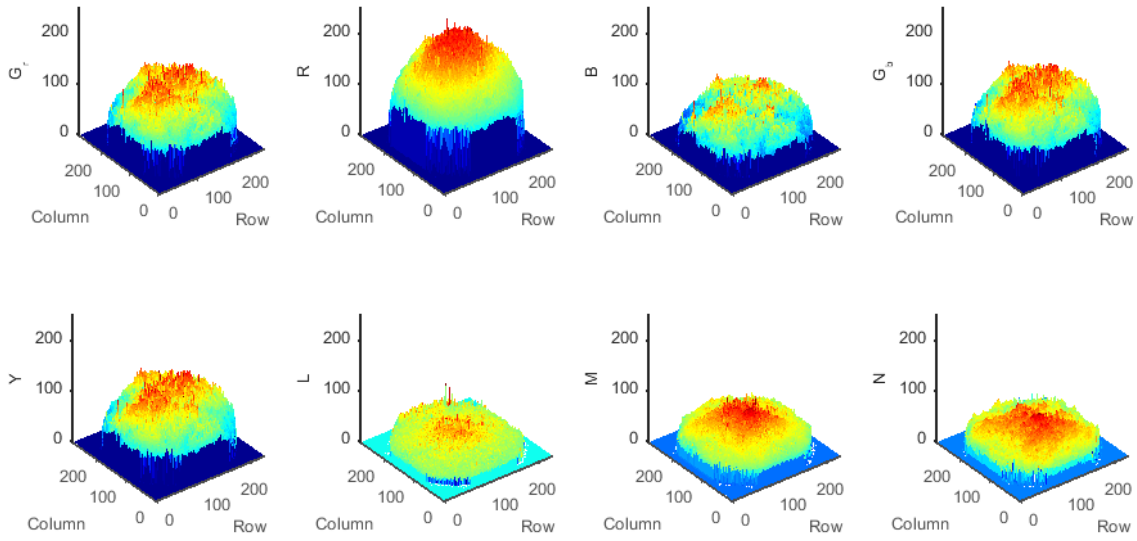


Figure 3-5. Surface plot of the colour components: Original GRBG (top row) and transformed YLMN (bottom).

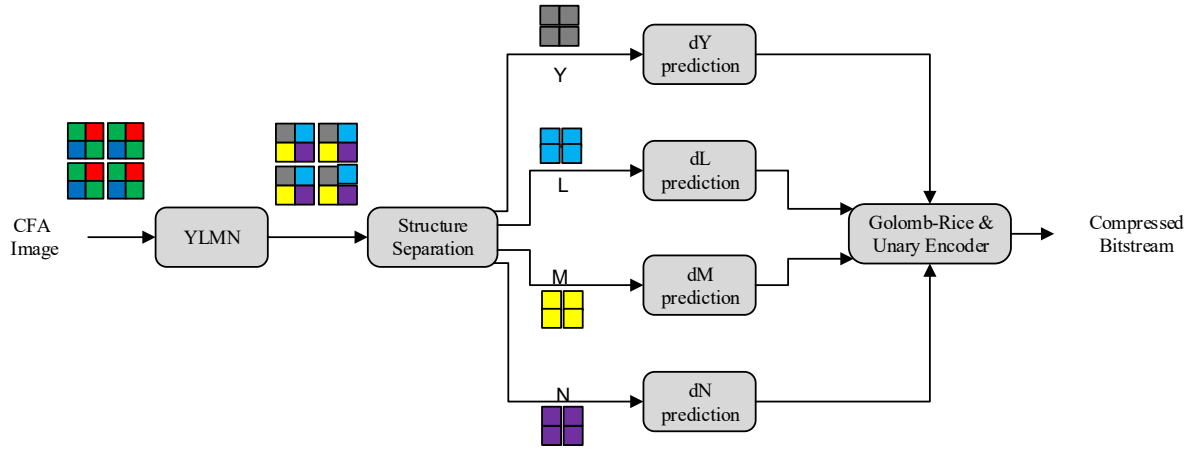


Figure 3-6. Block Diagram of the Proposed Lossless Compression Algorithm

3.3.1 Structure Separation

After the YLMN colour transformation, the proposed scheme deinterleaves the colour components into four sub-images. As previously mentioned in Section 3.2, direct compression of CFA image is not efficient as the mosaic arrangement of the colour pixel produces artificial high frequencies. By deinterleaving the CFA image, four downsampled sub-images, each of which consists of pixels in the single colour channel, can be extracted. This deinterleaved procedure removes the artificial high frequencies and therefore improves the compression performance. In the previous study in image compression for CE images [25], structure separation was achieved by shifting the G components to the left of the frame and shifting the R and B components to the right of the frame as shown in Figure 3-7(a). However, due to the requirement of buffer memory, this arrangement is not suitable for hardware implementation. Therefore, instead of rearranging the components in original CFA image, the prediction model is modified to generate prediction from same colour components as shown in Figure 3-7(b). This reduces the frame memory requirement and leads to a low-memory implementation in VLSI.

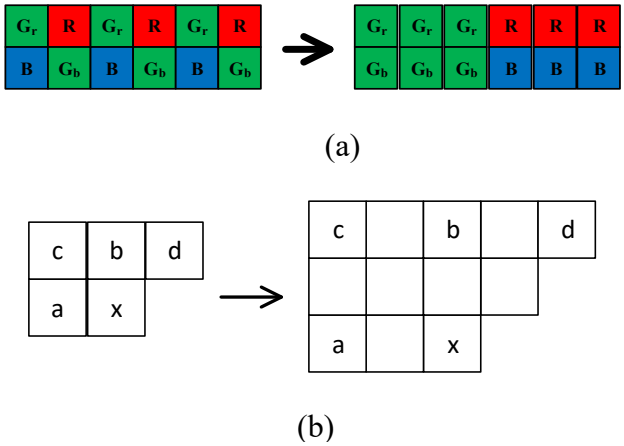


Figure 3-7. Different structure separation methods : (a) Structure Separation using rearrangement of the pixels, (b) Structure separation achieved by modifying the prediction template

3.3.2 DPCM

The compression efficiency of predictive coding depends on how accurate the prediction model can predict. Simple linear predictor often generates significant error around the edge areas. To achieve high prediction performance in the edge area, high complexity techniques such as template matching [62], hierarchical prediction [63] and context matching [47] has been employed in lossless CFA compression. However, these techniques require a high amount of frame memory and computational complexity. On the other hand, CE images generally have a larger smooth area than natural images [21]. Therefore, in comparison to simple prediction model, the improvement of coding performance by high-complexity prediction technique is not significant enough to justify the increase in compression overhead. Simple prediction model such as delta pulse coded modulation (DPCM) prediction model has shown impressive performance in terms of computational complexity and lossless compression ratio in capsule endoscopic image [64][59]. The proposed prediction model achieves a high prediction performance by exploiting the smoothness of the endoscopic CFA image after the YLMN colour transformation. It utilizes the horizontal prediction to handle the pixel in a conventional raster scan order without any buffer memory. In order to address the discontinuity between the neighboring pixels in CFA images, the predicted value is generated from nearest pixel from the same colour plane in the

horizontal direction. If the CFA image after YLMN transformation is denoted as X, then the residual prediction error is measured in the following way:-

$$dX(i,j) = X(i,j) - X(i,j-2) \tag{3-5}$$

As the application of YLMN transformation takes into account the inter-colour correlation, independent encoding of the deinterleaved sub-images using eqn. (3-5) yields near optimal compression performance for CE images.

3.3.3 Golomb-Rice Encoder

This section presents the single context Golomb-Rice encoder for the proposed compression system. In designing a prediction encoder, it is important to reduce the energy of the residual signal as much as possible. The YLMN transformation and the DPCM prediction model accomplish this goal by reducing the intra-spectral and inter-spectral redundancy. In addition, if the probability distribution function (pdf) of the residual signal can be correctly estimated, further reduction of the entropy can be achieved. Therefore, a single context is used for each sub-images, to estimate the pdf of a residual signal continuously while coding the residual signal.

The principal of the Adaptive Golomb-Rice coder used here is similar to JPEG-LS encoding [14]. However, instead of using 365 contexts for each colour plane, the proposed encoder only utilizes one context for each colour plane. This modification results in a drastic reduction of computational complexity and memory. The algorithmic description is given in Figure 3-8.

As the different colour components in the YLMN colour space have different characteristics and statistical properties, they are encoded separately with separate context. First, the residual signal is converted to a non-negative number using a mapping function as shown in the algorithm. Then the Golomb-Rice coding [65] represents the number into two strings. The first strings is an unary representation of the quotient $\left\lfloor \frac{M}{2^k} \right\rfloor$ as the prefix, while the second string is the fixed length code of the remainder. The length for coding residual error dX is $\left\lfloor \frac{2|dX|}{2^k} \right\rfloor + k + 1$.

Algorithm:- Encoding of Residual Error

Inputs: dY_{ij} : Prediction Error

Outputs: S_Y : Bitstream for Y

%% Mapping to Non-Negative Integers

if $dY_{ij} \geq 0$ **Then**

$$M \leftarrow 2 \times dY_{ij}$$

else

$$M \leftarrow 2 \times abs(dY_{ij}) - 1$$

end

%% Parameter k Estimation for Golomb-Rice

$k \leftarrow 0$

while $N_c \times 2^k \geq A_c$ **do**

$$k \leftarrow k + 1$$

end

%% Golomb Rice Coding to convert into variable length bitstream

$S_Y \leftarrow Golomb\_Rice(M, k)$

%%Update Context Parameter N_c and A_c

$N_c \leftarrow N_c + 1$

$A_c \leftarrow A_c + abs(dY_{ij})$

if($N_c > N_{Threshold}$) **Then**

$$N_c \leftarrow \frac{N_c}{2}$$

$$A_c \leftarrow \frac{A_c}{2}$$

end

Figure 3-8. Pseudo-code for residual signal encoding. It shows the encoding of the Y components. The encoding of L, M, and N also follows a similar procedure with different N_c and A_c values to store their context. Here, the function returns the bitstream to code the value M with parameter k using a Golomb-Rice entropy coder.

The parameter k is continuously updated based on the statistics of the signal using two registers to count the number of occurrences (N_c) and accumulate the error (A_c). In order to keep the cost of implementation low, a threshold value $N_{Threshold}=8$ was used. If N_c becomes greater than this threshold value, both N_c and A_c were halved.

3.3.4 Corner Clipping

The corner clipping mechanism presented in [17] is utilized in this work to discard the dark corner regions in the CE images. This allows to code the corner regions without any bits. As the size of the image and the shape of the corner region is known, the decoder can reconstruct the image using the bitstream. Therefore, this is an essentially a lossless mechanism.

3.4 Performance Analysis

In this section, the performance of the proposed image compression algorithm is assessed and compared with other lossless compression method developed for capsule endoscopy compression as well as lossless compression method such as JPEG-LS and JPEG-2000. Experiments are carried out using 100 images taken from KID Database [9]. 20 example images are shown in Figure 3-9. This database is chosen because it is a publicly available and contains capsule endoscopic images taken from a wide variety of location from the human gastrointestinal tract. All the images in the database are captured using Mirocam and stored in full resolution of 360x360 [42]. The full-colour RGB images in the database are sampled by the Bayer CFA pattern to produce the simulated CFA grayscale image. Performances of different lossless methods are evaluated by comparing compression rate (CR). The CR is reported in bits per pixel (bpp), which is defined as the ratio of the output bitstream in bits and the total number of pixels in the image.

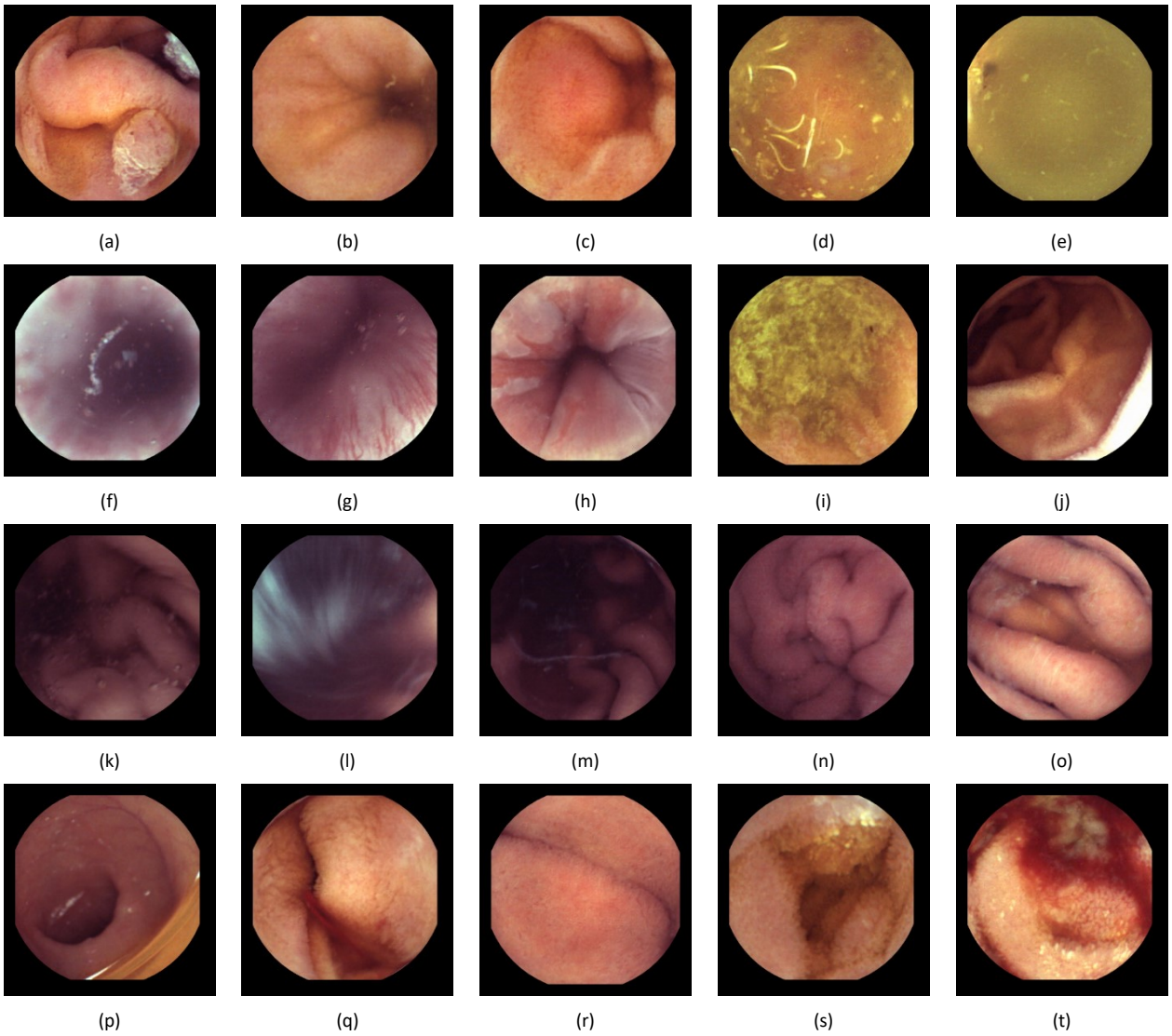


Figure 3-9. Example images from KID Dataset : 20 (out of 100) test images taken from [9]

3.4.1 Comparison of the GRBG and YLMN Colour Space

This section compares the compression performance of the original GRBG channels and the proposed YLMN colour space. For the first 5 test images in the database, Figure 3-10 shows bar chart of the standard deviation of each sub-images in GRBG colour space and the YLMN colour space. The height of each bar represents the standard deviation of the corresponding sub-image.

The figure shows that YLMN colour transformation significantly reduces the standard deviation of the sub-images, particularly for L, M and N channels. As lower standard deviation indicates smooth image, it is expected that L, M, and N-channel will result in a lower compression bitrate. As shown in

Table 3-1, the lossless compression rate is reduced particularly for L and N channels.

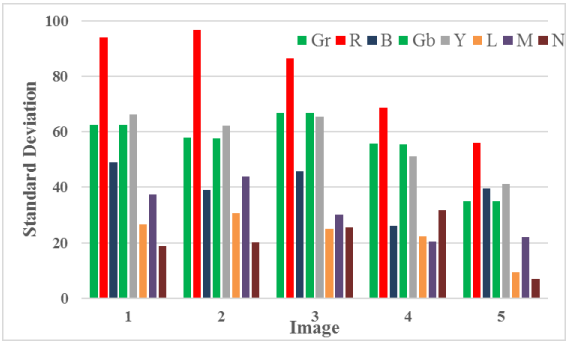


Figure 3-10. Comparison of standard deviation between the colour components in original colour space and YLMN colour space.

Table 3-1: Lossless bitrate of different color components (using the proposed compression scheme)

Image	Gr	R	B	Gb	Y	L	M	N
1	4.33	4.34	4.26	4.32	4.21	3.84	4.31	4.15
2	4.08	4.21	3.98	4.05	3.95	3.70	4.16	3.97
3	4.00	4.04	3.88	3.99	3.87	3.57	3.98	3.91
4	4.31	4.38	4.12	4.25	4.14	3.80	4.23	4.21
5	3.75	3.96	3.83	3.72	3.66	3.44	3.89	3.74

3.4.2 Comparison of Various Prediction Model

In this work, the prediction model is chosen based on the trade-off between computational complexity and prediction accuracy. In order to investigate the influence of the prediction model on coding performance, two median edge prediction (MEP) model is examined in this section.

First median edge predictor, denoted as MEP1, is widely used as a part of the JPEG-LS compression engine [14]. MEP1 consists of a flat region detector along with a causal template based median edge detection circuit to predict the current pixel from the neighboring pixels. In addition, the median edge prediction model presented in [11], which is denoted as MEP2, is also assessed. MEP2 discards the flat region detection to reduce the computational complexity. Both these two model is applied on the original CFA image where the context model is modified according to Figure 3-7. The proposed prediction model is applied on the original CFA image and on the YLMN image. Figure 3-11 shows the performance in terms of prediction gain and entropy of the residual signal.

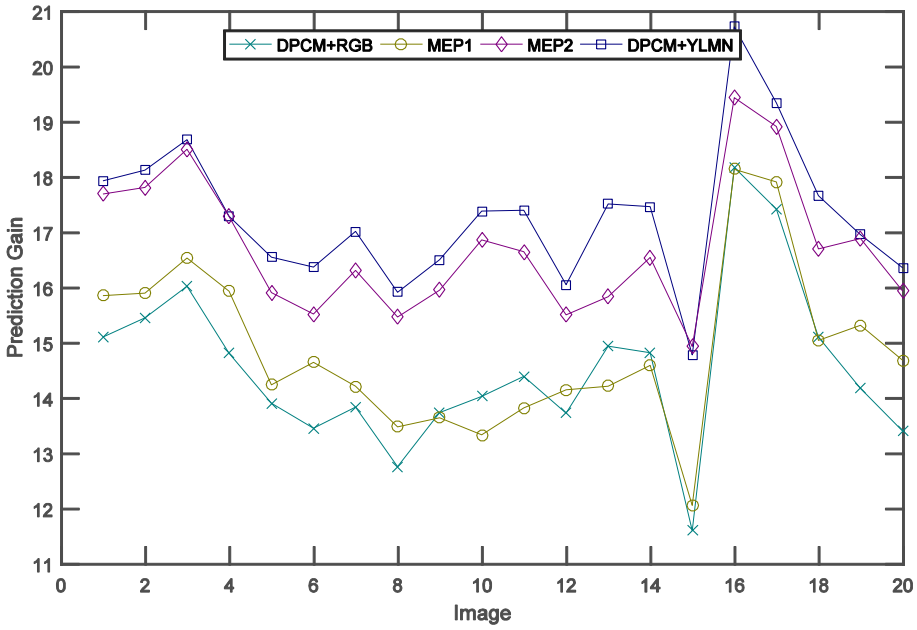
The prediction gain is defined as the ratio between the variance of the original signal to the residual signal measured as shown in eqn (3-6):-

$$G_p = 10 \log \frac{\sigma_{ori}}{\sigma_{pred}} \tag{3-6}$$

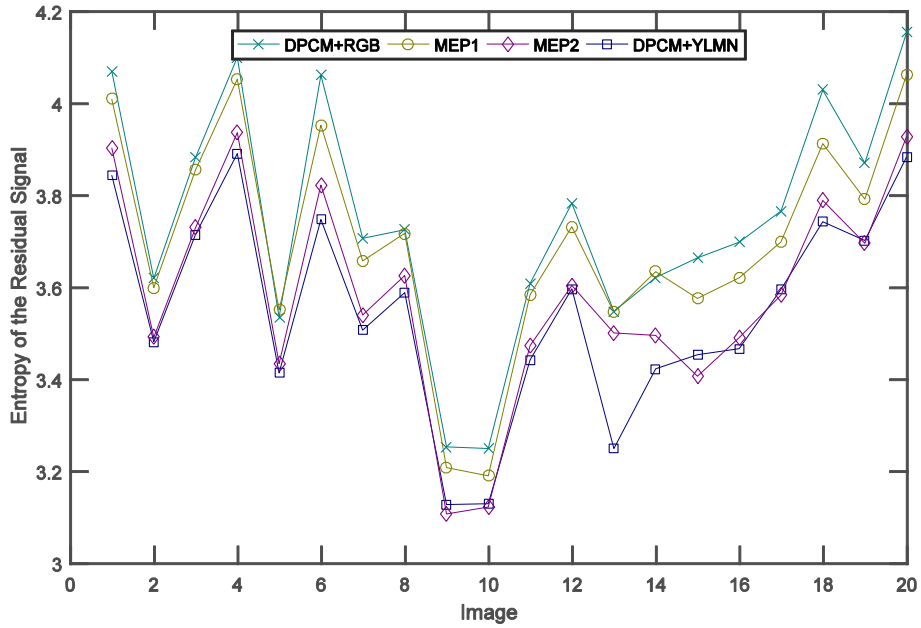
where, σ_{ori} and σ_{pred} are the variance of the original signal and residual signal respectively. On the other hand, the entropy of the image can be determined by eqn (3-7):-

$$H = -\sum_{i=1}^{n} P_i \log_2 P_i \tag{3-7}$$

where P_i is the probability of occurrence of intensity value i. Since the entropy of image date dictates the theoretical lower bound of lossless compression rate achievable, it assesses the efficiency of different prediction models. The proposed method with the YLMN colour transformation demonstrates the highest prediction gain and the lowest entropy, indicating a potential high compression efficiency. MEP2 can give a comparable result to the proposed method. However, as shown in Figure 3-7, the prediction model in MEP2 requires storing two line of data for the prediction, while DPCM needs no buffer memory. Therefore, the prediction model is expected to have lower memory requirement.



(a)



(b)

Figure 3-11. Comparison of (a) Prediction Gain and (b) Entropy of sample images from the KID database with various prediction model

Table 3-2 lists the computational complexity in terms of normalized operations, such as addition (ADD), shift (SHF), comparison (CMP). The buffer memory (MEM) requirement is also considered for the colour space transformation and prediction computation. As expected, DPCM with horizontal prediction provides the lowest complexity and memory requirement. However, Figure 3-11 (b) shows the entropy is highest for DPCM model as the simple model cannot exploit the inter-colour correlation sufficiently. Both MEP1 and MEP2 offers better compression performance with an increase in computational complexity and buffer memory. However, by employing the YLMN colour transformation, the proposed method smoothens the L, M and N channels. As a result, it can give the best entropy and prediction gain. The increase in the computational complexity and buffer memory due to YLMN colour transformation can be considered tolerable given that the proposed method yields reduction in average lossless bitrate of 0.2bpp as shown in Figure 3-11 (b).

Table 3-2: Number of operation per pixel required for various prediction model

Prediction Model	ADD	SHF	CMP	Total	MEM
MEP1	2	0	7	9	2 x Width
MEP2	3	1	4	8	2 x Width
DPCM	1	0	0	1	0
Proposed Method	2.5	0	0	2.5	½ x Width

3.4.3 Compression Performance

This section compares the coding performance of proposed algorithm with standard lossless engine such as JPEG-2000 (J2K) and JPEG-LS (JLS), and also with recently proposed lossless compression algorithms for CE: cost-efficient lossless compression (CELC) [28], computationally efficient image compressor (CEIC) [11], lossless image compression system (LICS) [10], and lossless low-power image compressor (LLIC) [29]. As JPEG-2000, JPEG-LS, LICS, LLIC and CEIC works on the full-colour image, bilinear interpolation is applied to interpolate the full-colour image from the original CFA image. Then the full-colour image is fed to these compression schemes, and the compression rate is measured by considering the size of original CFA image. Both CELC and proposed method works directly on raw CFA image, therefore demosaicking is not used in these cases.

Table 3-3: Lossless bitrate of various lossless compression scheme for CE

Image	JPEG 2000 [66]	JPEG LS [14]	Khan 2014 [59]	Fante 2015 [19]	Chen 2015 [32]	Proposed		
						DPCM +GR	YLMN+ DPCM+GR	YLMN+DPCM +GR+Clipping
1	7.01	5.71	7.07	9.98	4.15	4.31	**4.13**	**3.87**
2	6.50	5.21	6.73	9.44	**3.90**	4.08	3.95	**3.69**
3	6.18	4.95	6.54	9.03	**3.72**	3.98	3.83	**3.56**
4	6.93	5.67	7.14	9.72	4.13	4.27	**4.10**	**3.85**
5	6.06	4.82	6.22	8.64	**3.67**	3.81	3.68	**3.43**
6	6.71	5.42	6.96	9.64	4.07	4.28	**4.03**	**3.77**
7	6.31	5.03	6.64	9.19	3.88	4.06	**3.86**	**3.59**
8	6.67	5.36	6.76	9.49	3.96	4.08	**3.94**	**3.68**
9	5.62	4.52	6.17	8.16	**3.46**	3.68	3.53	**3.27**
10	4.91	3.94	5.67	7.18	**3.02**	3.33	3.15	**2.89**
11	5.33	4.30	5.92	7.84	**3.20**	3.47	3.32	**3.07**
12	5.06	4.08	5.65	7.29	**3.00**	3.21	3.12	**2.88**
13	6.00	4.82	6.19	8.68	**3.57**	3.79	3.68	**3.41**
14	6.56	5.28	6.67	9.45	3.94	4.11	**3.93**	**3.67**
15	4.78	3.87	5.61	6.85	**2.74**	2.89	2.88	**2.65**
16	5.63	4.49	6.07	8.21	**3.35**	3.70	3.43	**3.28**
17	6.33	5.08	6.62	9.26	3.84	4.09	**3.82**	**3.66**
18	6.68	5.40	6.96	9.68	4.10	4.35	**4.09**	**3.81**
19	6.72	5.43	7.12	9.97	4.06	4.36	**4.04**	**3.88**
20	7.07	5.71	7.27	10.11	4.20	4.41	**4.18**	**3.92**
Avg	6.25	5.05	6.58	8.99	3.75	3.95	**3.73**	**3.53**

Except for JPEG-2000, all the other methods employ a predictive coding technique for lossless encoding. In LLIC, an RGB demosaicked image is first mapped to YUV colour space. Then, a static DPCM prediction model is used for computing the residual signal, which is encoded by a Golomb-Rice encoder. A clipping scheme is also employed for removing the

corner regions. LICS improved upon LLIC by proposing a dedicated colour space YEF, which can capture the unique characteristics of endoscopic image compression. YEF colour transformation reduces the complexity of colour transformation as well as significantly improves the coding performance of the chrominance channel. Note should be taken that both YUV and YEF colour transformation is not integer reversible and requires up to 3 additional fraction bits for lossless colour transformation. The lossless compression scheme in CEIC consists of a reversible colour transformation YrUrVr, DPCM, and corner clipping. In CELC, the framework of JPEG-LS is modified for designing a cost-efficient image compressor for raw CFA image based on MEP2. The context template in the prediction model was modified to take account of the interleaved pixels in the raw CFA image. A hybrid encoder consists of Huffman and Golomb-Rice coder is used for efficient encoding of the residual error.

The output compression bitrates for CE images from KID Database achieved by various methods is listed in Table 3-3. The results clearly illustrate that the algorithms based on a full-colour image are not efficient for lossless encoding of raw CFA image. This indicates the compression stage cannot completely remove the redundancy added by the demosaicking stage. Among these methods, JPEG-LS gives the best result in terms of compression efficiency, outperforming second best LICS by 0.7 bpp.

On average, the proposed algorithm, with YLMN colour transformation, Golomb-Rice coding, and clipping, can outperform all the other methods by achieving the lowest lossless compression of 3.53 bpp. Even without clipping, the proposed method has comparable results with CELC, which has a higher computational complexity as shown in Figure 3-11. Comparison of (a) Prediction Gain and (b) Entropy of sample images from the KID database with various prediction model

Table 3-4 shows the comparison of the proposed method with other existing works on image compression for CE. Table 3-4 also lists different feature of these algorithms such as: whether the algorithm uses demosaicking-first compression-later (DC) scheme or compression-first demosaicking-later scheme (CD), the colour transformation (CT) used and the core compression algorithm used. In order to distinguish between the colour transformation for full-colour image and CFA image, three components are used for full-colour image, and four

components are used for CFA colour transformation. For example, the original colour space in full-colour images is expressed as RGB, while the companion colour transformation in CFA is expressed as GRBG. From the observed results, it is evident proposed algorithm can give the lowest bitrate among all the other lossless compression algorithm. The lossless bitrate of proposed algorithm is greater than or comparable to the near-lossless (PSNR>46dB) bitrate achieved by [67], [19], [25] and [20].

Table 3-4: Comparison with other compression schemes

Type	Work	DC / CD	CT	Algorithm	BPP	PSNR
Lossy	Wahid et. al. [26]	CD	GRBG	DCT	1.03	32.9
	Turzca et. al. [68]	CD	YCoCgY	DCT	0.77	36.5
	Lin et. al. [69]	CD	GRBG	DCT	1.63	32.5
	Dung et. al. [28]	CD	GRBG	DCT	1.44	36.2
	Li et. al. [67]	DC	RGB	DCT	5.91	47.7
	Lin et. al. [70]	CD	GRBG	DCT	1.42	40.7
	Chen et. al. [25]	CD	GRBG	Pred.	3.46	46.4
	Turzca et. al. [12]	CD	YCuCvY	DCT	0.70	35.7
	Liu et. al. [20]	CD	GRBG	Pred.	3.46	46.3
	Fante et al. [19]	DC	YUV	Pred.	6.33	40.6
Lossless	Khan et al. [17]	DC	YUV	LPC	6.48	∞
	Khan et al. [59]	DC	YEF	Pred.	5.28	∞
	Fante et al. [19]	DC	YUV	Pred.	6.33	∞
	Chen et al. [32]	CD	GRBG	Pred.	5.49	∞
	Proposed	CD	YLMN	Pred.	**3.53**	∞

Apart from the lossless bitrate, the computational complexity of the proposed compression method is also assessed. To this end, this work was realized by using a hardware description language (HDL) Verilog and an EDA tool Design Compiler was used to synthesize the proposed design by TSMC 65 n*m* CMOS process. The gate counts of this work are only 3.78 k, and the core area is 29,627.28 (um x um). The power consumption of this work is 0.9 mW simulated by using SYNOPSYS Design Compiler when it operated at 250 MHz. The ASIC implementation of the design is built to check the utilization of resources and compare with the other algorithm. The resolution of the image used in this implementation is 640 x 640. Table 3-5 shows the comparison with other lossless and near-lossless encoder for CE. This implementation reduces the gate counts by 38.9% and memory requirement by 71.2% than the previous design. The underlying reasons can be summarize as: (i) application of low complexity DPCM prediction model with a low complexity YLMN colour transformation to exploit both spectral and spatial redundancy, (ii) usage of a memory efficient Golomb-Rice coder instead of a Huffman Coder, (iii) application of a low complexity corner clipping mechanism to cut the uninformative regions in the corner.

Table 3-5: Comparison with other lossless encoders for CE[1]

	[21]	[22]	[19]	[14]	[18]	[23]	This Work
Year	2009	2007	2009	2007	2010	2014	2016
Process(um)	0.180	0.180	FPGA	0.180	0.180	0.180	0.065
Frequency	20~24	20	21	183	200	200	250
Resolution	640	640	640	640	640	640	640
Gate Counts	19.5	50	37.3	27.68	11.57	5.54	3.78

[1] The proposed algorithm was implemented in Hardware description language by KMM Rahman

Memory(k)	17.5	29	146	20.81	15.2	10.2	2.93
Power(mW)	1.3	3.5	NA	NA	4.1	2.2	0.94
Normalized area	5.16	13.23	9.86	7.30	3.06	1.46	1

Note: The normalized core area is normalized by the NAND-equivalent gate counts.

3.5 Conclusion

In this chapter, a novel lossless image compression scheme is proposed for wireless capsule endoscopy system. The proposed system utilizes an optimum reversible colour transformation to reduce the spectral redundancy in the image and adopts a low memory raster order DPCM prediction model to reduce the spatial redundancy in the image. Generated residual signal are then compressed by an adaptive Golomb-Rice encoder with a single context, that effectively encode the signals with very low computational complexity and memory requirement. A low complexity corner clipper scheme reduces the lossless bitrate by removing the uninformative corner regions in the image. Experimental results show that the proposed algorithm outperforms other lossless compression algorithms in terms of lossless compression rate, gate counts, and memory requirement.

Chapter 4 - Low-Bitrate and High-Quality Image Compressor for Wireless Capsule Endoscopy

The chapter includes a manuscript entitled ‘Low-Bitrate and High-Quality Image Compressor for Wireless Capsule Endoscopy’ by Shahed K. Mohammed and Khan A. Wahid. The previous chapter proposed a lossless compression algorithm. The algorithm is able to transmit images in a lossless manner within the bandwidth requirement for a frame rate of 6 frames per second. This chapter presents a lossy compression algorithm that enables high-quality image transmission with higher frame rate, which is a necessary condition for integrating accurate remote manipulation and therapeutic maneuvers in the WCE.

Abstract—This chapter presents a dedicated image compressor for wireless capsule endoscopy system that works directly on raw colour filter array (CFA) images. The proposed method involves a novel low complexity YEFD colour transformation to decorrelate the colour components in CFA image, a structural separation to reduce the artificial high frequencies between the pixels, a low complexity DCT like image transformation and visually optimal quantization matrix to reduce the spatial redundancy, and run-length coder along with adaptive Golomb-Rice entropy coder to efficiently encode transformed coefficients. A fast and high-performance demosaicking algorithm is presented to reproduce the colour image from the reconstructed CFA image. This algorithm takes into account the error generated by the lossy compression algorithm to reduce the error propagation in the demosaicking stage. The experimental results demonstrated that the proposed algorithm could achieve compression ratio up to 12 with peak signal to noise ratio as high as 40dB. The proposed algorithm can outperform all the other state-of-the-art image compression algorithm in terms of compression ratio and image quality.

***Index Terms*—Colour filter array (CFA), colour transformation, demosaicking, lossy compression.**

4.1 Introduction

In a small timeframe of 15 years, wireless capsule endoscopy (WCE) has created a paradigm shift in gastrointestinal (GI) disease management [1]. Initially launched as a tool to visualize the small-bowel, WCE has now become a patient friendly non-invasive alternative to screen the entire GI tract [2]. The pill size device consists of a relatively low-resolution image sensor, an LED-based illumination system, a data processing unit, a battery and a radio transmitter. After being swallowed by the patient, the capsule travels GI tract using the peristalsis motion, captures images and transmits them wirelessly to the recorder outside throughout the journey which can span up to 8 hours [71]. Current research efforts are focusing on incorporating functionality to obtain the biopsy and perform therapeutic maneuvers, implementing automatic detection of regions of interest to reduce the reading time, and introducing active locomotion to control the capsule movement [2], [42], [71]. Accurate remote manipulation required for therapeutic intervention and obtaining biopsy necessitates a higher frame rate and image

resolution. Therefore, the improvement of image resolution and frame rate while keeping the battery size low plays a pivotal role in the development of next generation capsule.

The image quality in terms of resolution and frame rate is limited by the power consumption due to wireless communication in the capsule [10]. Moreover, the available bandwidth for transmitting data through the human body is restricted to 2~3 Mbps by Federal Communications Commission [11]. Therefore, the compression of the data before transmission is essential for improving the image quality and increasing the battery life. Keeping these constraints in mind, the compression rate required to send 512x512 images in 8-bit raw CFA format should be between 6 to 15 depending on the frame rate. Such high compression ratio is not feasible in lossless compression. So currently, lossy compression is preferred in WCE [42]. On the other hand, as these images contain medical data, the image should be transmitted with little or no distortion so that the clinically relevant information are preserved to aid both manual and automated diagnosis.

Most of the image sensors capture the frame using a single sensor array with a Bayer colour filter array (CFA) placed in front of each pixel to capture different colour in different pixel [56]. As standard compression algorithms such as JPEG, JPEG 2000, MPEG or H.264 are designed for full-colour image, a demosaicking stage interpolates the missing colour information before compression. This framework is defined as the demosaic-first scheme, where the compression occurs on the full-colour image. Two major issues arise due to the use of a demosaic-first scheme in WCE. First, the low complexity demosaicking stage introduces significant colour artifacts particularly in the region at the boundaries, that will cause distortion of clinically relevant findings. By using an appropriate edge preserving algorithm, the image quality can be significantly improved, but it will increase the computational complexity of the compressor [41]. Second, from a data compression perspective, the demosaicking stage adds data without any new information leading to a sub-optimum compression performance. On the other hand, compressing the raw CFA image and transmitting the image to the receiver allows the utilization of effective post-processing methods such as edge preserving interpolation and denoising. Therefore, a compression-first lossy algorithm is ideal for WCE that will reduce the computational complexity as well as increase the compression rate for the endoscopic system [12], [26], [28], [32], [38], [39], [69], [72], [73].

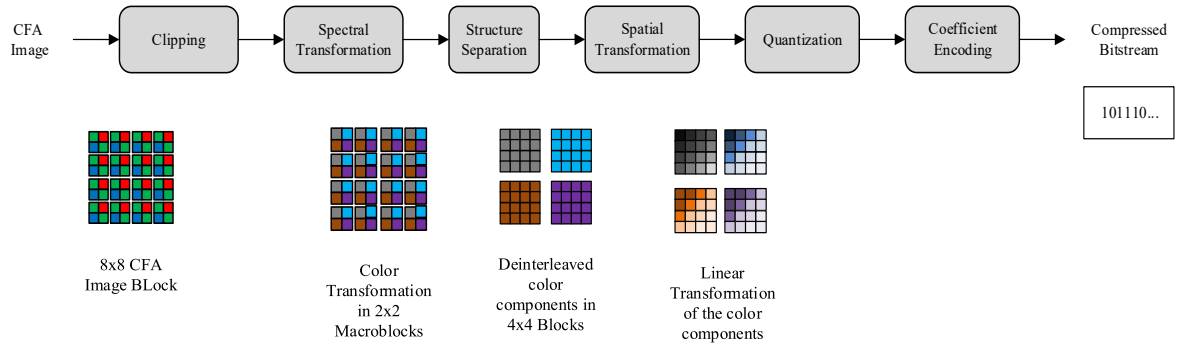
Although there are few low-complexity prediction-based systems, their compression ratio is sub-optimum for WCE. Therefore, most of the lossy algorithms [12], [26], [28], [38], [39], [69] on WCE employ transform coding based on discrete cosine transform (DCT) similar to JPEG and H.264 compression engine. DCT has a very high energy packing efficiency which allows the image compressor to represent the image with a few coefficients with performance comparable to optimal transform coding such as Karhunen Loeve Transform (KLT). However, these methods transform the image in original colour space (RGB) or use colour transformation optimum for natural images rather than endoscopic image. With the distinct red hue, endoscopic images occupy a narrow band in the natural image. A compression scheme adapted based on natural image colour transformation cannot sufficiently exploit the inter-colour correlation and characteristics in endoscopic images.

This chapter presents a new image compression algorithm for WCE. This work focuses on raw endoscopic images sampled with Bayer CFA pattern due to its popularity in the industry. The proposed method consists of colour transformation, colour channel separation, integer discrete cosine transformation, visually optimum quantization and adaptive Golomb-Rice (AGR) coding. The decoding circuit includes inverse transform and a modified demosaicking method to increase the image quality.

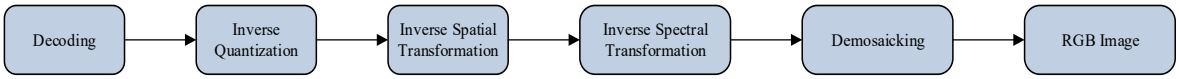
The contributions of this work are listed below. First, this work introduces a new colour space transformation optimum for endoscopic image compression. The proposed colour space transformation operates on the 2x2 macroblock in CFA image to reduce the redundancy of the colour components and separate the structure for efficient compression of the mosaic image. Besides, the proposed method packs the majority of variance into one plane and therefore enables the unequal bit allocation by the application of quantization in the following stages. Secondly, a low complexity DCT like transformation is employed here which can effectively pack the energy in individual colour components. For the transformation, rate-distortion optimized quantization matrices are derived from the endoscopic images. The coefficients are encoded using a low-complexity context adaptive Golomb-Rice run-length encoder. A modified demosaicking algorithm is proposed to denoise and demosaic the colour image simultaneously.

This chapter is organized as follows. First, section 4.2 provides a brief overview of the proposed system. In section 4.2.1-4.2.5 outlines the image compressor algorithm. The

performance of the image compressor and the results of its comparison to JPEG and JPEG-2000 are described in section 4.3. Section 4.4 contains the conclusion.



(a)



(b)

Figure 4-1. Block Diagram of the (a) Proposed Image Encoder and (b) Decoder

4.2 Proposed Algorithm

Figure 4-1 illustrates the proposed image encoding and decoding method for WCE. In CFA image, adjacent pixels come from different colour components. So the resultant interleaved image contains significant amounts of energy in the high-frequency spectrum. Therefore, direct compression of CFA image cannot achieve high compression ratio and produce repeating distortion pattern in the reconstructed image [17]. To overcome this issue, this work employed structure separation by deinterleaving and downsampling the colour components in the CFA image. The proposed system also employs a colour transformation to map each 2x2 Bayer cell in original CFA to one luminance channel and three chrominance channels. The transformed components have lower inter-spectral redundancy than that of original components. After the colour transformation, this work employed a structure separation to extract four sub-images from the original CFA image. Then each sub-image is processed by block-based transform coding to efficiently pack the energy of each block in a few coefficients and reduce the spatial redundancy in each sub-images. This work has chosen the optimum block-based transform through

experiments on energy packing efficiency and computational cost. An integer approximation of the optimum transformation is employed to reduce the computational complexity. The transformed coefficients are then quantized to achieve optimum rate-distortion performance as well as to avoid loss of clinically relevant findings. Then the quantized AC coefficients are processed using run-length encoder along with a low complexity adaptive Golomb-Rice encoder. On the other hand, the DC coefficients are encoded using delta pulse coded modulation (DPCM) followed by adaptive Golomb-Rice (AGR) encoder.

In the companion decoder (Fig. 1(b)), the generated sub-images undergo inverse colour transform to produce the original CFA layout. Finally, an edge preserving demosaicking algorithm reconstructs the full-colour image.

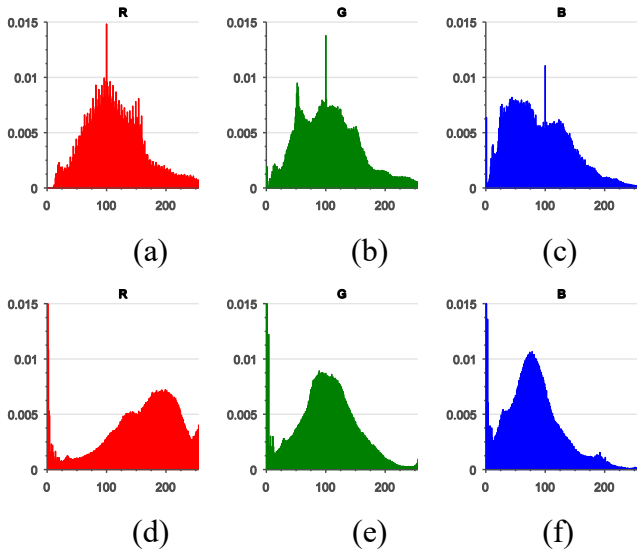


Figure 4-2. Comparison of histograms of the three colours : red(left), green (center) and blue (right) for Natural Images taken from Kodak Dataset (Above) and Capsule Endoscopy Images taken from Gastrolab (Below)

4.2.1 YEFD Colour Transformation:

This section explains the derivation of the proposed YEFD colour transformation. In CFA input data, individual sub-images generated by structure separation exhibits a high inter-spectral correlation similar to the colour components in the full-colour image. Therefore, individual encoding of the sub-images does not provide the best result as the correlation between

the colour components is not sufficiently exploited. In previous work on endoscopic compression, it has been shown that colour transformation can be utilized to exploit the unique characteristics of the endoscopic image, leading to superior compression system [64], [74]. To decorrelate the sub-images and capture the unique characteristics of capsule endoscopy (CE) image, a YEFD colour transformation method is proposed.

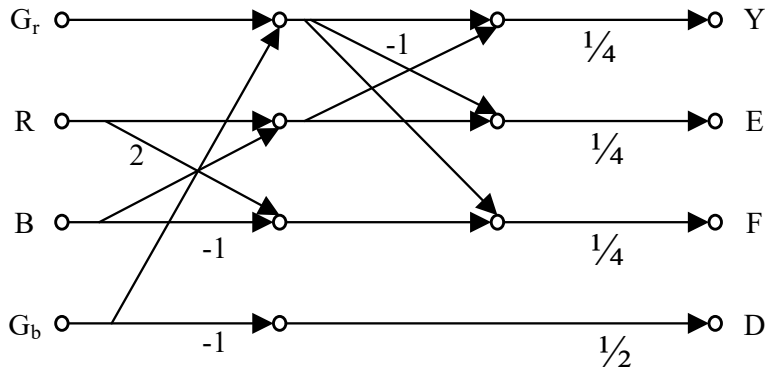


Figure 4-3. Implementation of the proposed colour transformation

To illustrate the distinctness of CE from natural images, the intensity distribution of 24 natural images taken from the popular Kodak Dataset [75] and 24 CE images from Gastrolab database [61] is compared. The CE images were taken from different locations in the gastrointestinal tract from esophagus to anus. For each image, the corresponding CFA image is simulated using a sampling pattern mimicking the Bayer pattern. Then using structure separation, sub-images for R, G, and B are generated. The aggregated histogram for each colour component is shown in Figure 4-2. As can be seen, the histograms in the CE has a much regular shape than the natural images. This is expected as the CE images only represent a narrow spectrum of the natural image. Therefore, to capture the unique characteristics of CE, principal component analysis (PCA) [76] is applied to derive the optimum colour transformation. For this purpose 200 images taken from 20 different location of the gastrointestinal tract is used [61]. The dataset contains normal finding images as well as disease condition such as cancer, polyp, coeliac disease, Crohn's disease, ulcer. The result is shown in eqn (4-1):

$$A = \begin{bmatrix} 0.2688 & 0.1752 & 0.2870 & 0.2689 \\ -0.0653 & 0.6158 & -0.2538 & -0.0652 \\ -0.2481 & 0.1106 & 0.3953 & -0.2460 \\ -0.4982 & -0.0005 & -0.0015 & 0.4998 \end{bmatrix} \tag{4-1}$$

where the colour transformation achieved by applying eqn (4-2)

$$\begin{bmatrix} Y \\ E \\ F \\ D \end{bmatrix} = A \begin{bmatrix} G_r \\ R \\ B \\ G_b \end{bmatrix} \tag{4-2}$$

As can be seen from eqn (4-1), the resultant transformation contains floating point operation. To avoid floating point operation, the approximated transform is obtained by rounding the scaled coefficients of the transform matrix. To select the optimum transformation, the local coding gain (LCG) measure [19] is used. LCG is defined as the average of the ratio of arithmetic mean to geometric mean of the variance of the transformed components in a NxN block of the image.

$$LCG = 10 \times \log_{10}\left(\frac{AM}{GM}\right) \tag{4-3}$$

Table 4-1: Comparison of local coding gain of different approximations (using 4x4 blocks)

Transformation	LCG	Transformation	LCG	Transformation	LCG
A	2.72	$\frac{1}{8}[A \times 8]$	**2.88**	$\frac{1}{64}[A \times 64]$	2.71
$\frac{1}{2}[A \times 2]$	0.90	$\frac{1}{16}[A \times 16]$	2.74	$\frac{1}{128}[A \times 128]$	2.71
$\frac{1}{4}[A \times 4]$	2.28	$\frac{1}{32}[A \times 32]$	2.72	Eqn (5)	**2.89**

The higher the gain, the higher compression is achieved in the transformation. The local coding gain is calculated in 4x4 blocks for each approximation as shown in Table 4-1. Here the [ ] indicates the rounding operation. It is evident that the local coding gain is maximum with a multiplication of 8, in which case, the eqn (4-1) can be written as:-

$$A = \frac{1}{8}\begin{bmatrix} 2 & 1 & 2 & 2 \\ -1 & 5 & -2 & -1 \\ -2 & 1 & 3 & -2 \\ -4 & 0 & 0 & 4 \end{bmatrix} \tag{4-4}$$

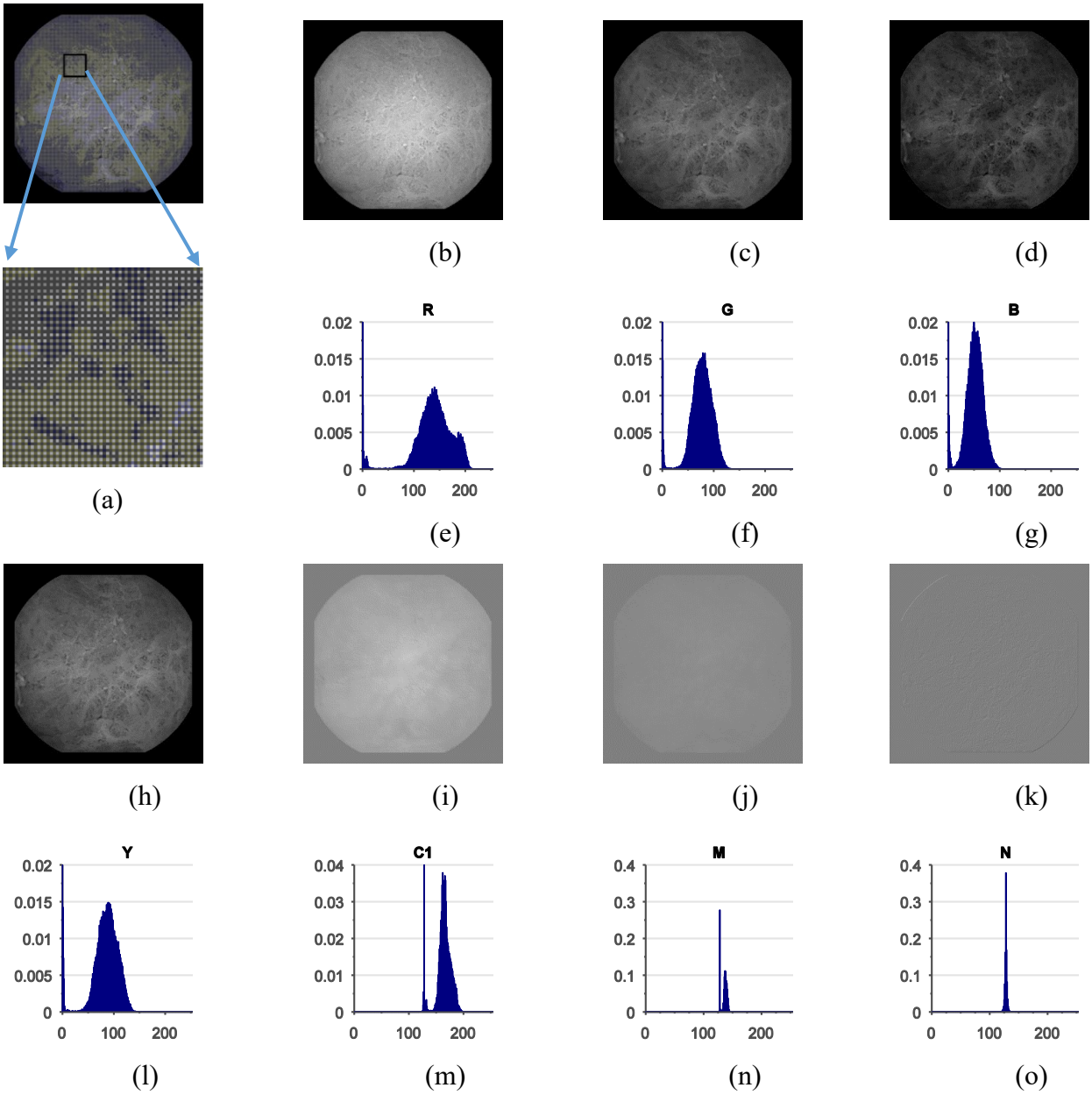


Figure 4-4. Comparison of sub-images in original and transformed colour components (a) Original Image in Bayer Format, with the bottom image showing the zoom of the image inside the black window. (b)-(d) Respectively the red, green and blue sub-images generated by separating the

colour components in the original image, while €-(g) are their respective histogram. The correlation between the sub-images is clearly visible. (h)-(k) the sub-images formed by separating the transformed colour components using the proposed colour space and (l)-(o) are their respective histogram. All the details are condensed in one sub-image Y, while the other components are very smooth as evident from both the image and histogram.

To ease the hardware implementation further, the following simplification as shown in eqn (4-5) is proposed. This simplified colour transformation can be implemented using only seven adders as depicted in Figure 4-3. Moreover, this approximation leads to increase in the local coding gain further as shown in Table I.

$$\begin{bmatrix} Y \\ E \\ F \\ D \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 2 & 2 & 2 & 2 \\ -1 & 4 & -2 & -1 \\ -2 & 2 & 2 & -2 \\ -4 & 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} G_r \\ R \\ B \\ G_b \end{bmatrix} \tag{4-5}$$

A visual comparison of the YEFD colour components and the original colour components is shown in Figure 4-4, which indicates that the proposed colour transformation yields good decorrelation capability through which most of the details of the image are concentrated in the luminance channel Y. On the other hand, the three chrominance channels contain less information and smoother surfaces, which are ideal for high compression.

The proposed YEFD colour space has several interesting properties. First, the luminance channel in the proposed colour space is a weighted sum of the four original values. The Y component holds the information for texture such as blood veins, mucosa structure. Disease condition such as angiodysplasia, chicken skin, celiac disease, Crohn's disease, polyp, and tumor are clearly discernible in this channel. The E channel is dominated by the red colour. Therefore underlying blood vessels and micro-vessels are better visible in this channel. F channel is the difference between the luminance and the average of the two green values. It has been shown in [74] that the luminance channel and the green channel has similar intensity distribution. Therefore, the resultant difference colour space has a very smooth surface with very small variance. The fourth channel (D) contains the difference between the two green channels. In a

smooth image, this channel will have a flat distribution and therefore a small entropy. It can be noted that the second channel is not required to reconstruct the green channel.

4.2.2 Structure Separation

After the YEFD colour transformation, the proposed scheme deinterleaves the colour components into four downsampled sub-images, each of which has only pixels from single colour components as shown in Figure 4-5.

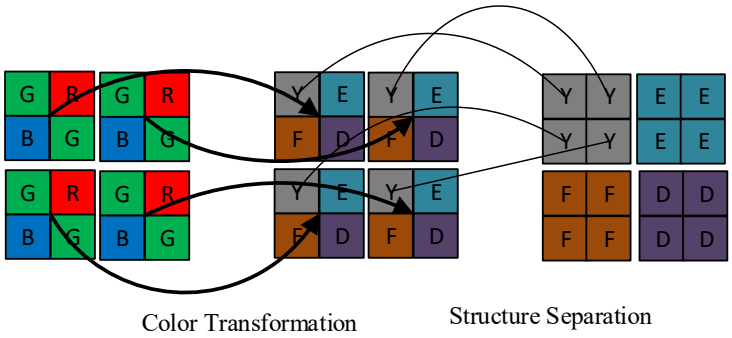


Figure 4-5: Bayer CFA colour transformation and structure separation method. The bold arrow represents the colour transformation of a 2x2 macroblock, while the thin line represents the structure separation procedure for Y sub-image. The other three sub-images can be generated in a similar manner.

Let us consider a $H \times W$ grayscale CFA image is represented by a matrix Z, where each 2×2 macroblock has been transformed using the YEFD colour space; then the structure separation step can be implemented using the following way:-

$$\begin{aligned} Y(i,j) &= Z(2i-1, 2j-1) \\ E(i,j) &= Z(2i-1, 2j) \\ F(i,j) &= Z(2i, 2j-1) \\ D(i,j) &= Z(2i, 2j) \end{aligned} \tag{4-6}$$

where $i = 1, 2, ..., \frac{H}{2}$ and $j = 1, 2, ..., \frac{W}{2}$. The obtained images are all of the quarter sizes of the original image. As each sub-image contains pixel from the same colour plane, the subsequent block-based transform can effectively exploit the spatial redundancy in the individual sub-images.

4.2.3 Image Transformation

The compression efficiency of transform coding depends on the ability of the transformation to pack energy into a few coefficients. Linear orthogonal transforms such as discrete cosine transform (DCT), discrete wavelet transform (DWT), discrete Walsh-Hadamard (DWHT) or discrete sine transform (DST) have demonstrated excellent decorrelation and energy compaction capability from a block of pixels. Particularly, DCT has shown to achieve comparable performance to the optimal transform, the Karhunen-Loeve transform (KLT), for natural images.

Table 4-2: Comparison of coding gain of different linear transformation in compressing different colour components

Transformation	Y	E	F	D	Complexity
KLT	9.96	7.92	5.27	1.52	96fa+192fm
DCT	9.94	7.97	5.32	1.42	72fa+8fm
DST	6.54	6.71	5.34	0.95	56fa+192fm
Hadamard	8.36	6.88	4.67	1.34	**64a**
H.264	**8.53**	**7.16**	**4.79**	1.53	**64a+16s**
IBT-4[21]	3.73	3.63	3.01	1.36	**48a**
HEVC	8.23	7.01	4.79	**1.54**	64m+56**a**

fm: float point multiplication, fa: float point addition, a: integer addition, m: integer multiplication, s: shift operation

Another factor in determining the optimum transformation is the block size. The buffer size required converting the progressive raster order data to block-wise data directly depends on the block size. In this work, the 4x4 block size is chosen for two reasons. First, due to the nature of CFA image, the 4x4 block in individual colour component represents an 8x8 block in the full-colour image. In previous work, it has been shown that the energy packing efficiency of 16x16 block is only marginally better than 8x8 block for full-colour image [12]. Therefore, 4x4 block size is chosen for CFA images as it gives a good trade-off between complexity and energy packing efficiency. Secondly, the ringing artifacts are more prominent in 8x8 blocks than 4x4 blocks as a result most recent video coding utilizes 4x4 image blocking such as H.264 [77]. To find the optimum 4x4 transformation for different colour space, the coding gain (G_{TC}) [78] of each transformation is calculated using eqn (4-7):-

$$G_{TC} = \left\{ \prod_{i=1}^{M} \left[\left(\frac{\sigma_{yi}^2}{\sigma_x^2} \right) \|f_i\|^2 \right] \right\}^{-(1/M)} \tag{4-7}$$

Where σ_x^2 is the variance of the input signals, σ_{yi}^2 is the variance of the i-th transformed coefficients and $\|f_i\|^2$ is the L2-norm of the i-th synthesis basis function of the transform matrix. The results are shown in Table 4-2. It is evident from the results, DCT has a comparable performance to the optimum transformation for all the colour components. Among the different integer approximation of DCT, IBT-4 [21] has the lowest complexity with lowest coding gain. It is evident the H.264 transformation has the highest coding gain with a comparatively low complexity.

Based on the experiments, the 4x4 H.264 transformation is selected for all the colour components. As described in [77], the 2D DCT of a 4x4 pixel blocks X can be computed as:-

$$X_{DCT} = (C_F X C_F^T) \circ \tag{4-8}$$

where $\circ$ represents the Hadamard product and T denote the transposition. Here C_F is the core transform of the integer approximation of 1-D DCT for forward transform, while S_F is a scaling matrix required for ensuring the orthogonality. To reduce the implementation cost, S_F is incorporated with the quantization matrix. The core transform used here is defined as below:

$$C_F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad S_F = \begin{bmatrix} \frac{1}{4} & \frac{1}{2\sqrt{2}} & \frac{1}{4} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2\sqrt{2}} & \frac{1}{4} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \end{bmatrix} \tag{4-9}$$

Once the transformed coefficients are obtained, they are quantized using an optimized quantization table using eqn. (4-10)

$$X_{Quantized} = X_{DCT} \varnothing Q \tag{4-10}$$

Where Q is quantization table and $\varnothing$ represents a Hadamard division. The goal is to achieve optimum rate distortion while keeping the implementation cost of scaling matrix and quantization matrix low. To this end, the scaling matrix is incorporated with the quantization matrix using the following formula:-

$$\hat{Q}_F = Q_F \varnothing S_F \tag{4-11}$$

Then quantization table $\hat{Q}_F$ is optimized based on the rate-distortion theory with a simple optimization algorithm [79]. The computational cost is minimized by forcing the quantization step to be a power of 2. First using the near-optimal algorithm presented in [79], the convex hull for rate-distortion profile is obtained. Then the bisection method is applied to find the quantization table to keep the peak signal to noise ratio greater than 40dB. The obtained quantization matrix for different colour components are shown below:-

$$\hat{Q}_y = \begin{bmatrix} 16 & 32 & 32 & 64 \\ 32 & 64 & 64 & 128 \\ 32 & 64 & 32 & 64 \\ 64 & 64 & 64 & 128 \end{bmatrix} \qquad \hat{Q}_E = \begin{bmatrix} 32 & 64 & 32 & 64 \\ 64 & 128 & 64 & 128 \\ 64 & 64 & 32 & 64 \\ 64 & 128 & 64 & 128 \end{bmatrix} \tag{4-12}$$

$$\hat{Q}_F = \begin{bmatrix} 32 & 64 & 64 & 64 \\ 64 & 128 & 32 & 64 \\ 32 & 64 & 32 & 64 \\ 64 & 64 & 64 & 64 \end{bmatrix} \qquad \hat{Q}_D = \begin{bmatrix} 32 & 64 & 32 & 64 \\ 64 & 128 & 64 & 128 \\ 32 & 64 & 32 & 64 \\ 64 & 64 & 64 & 128 \end{bmatrix}$$

After quantization, the transformed coefficients are zigzag scanned to transform the coefficients in the run-length format following the JPEG baseline standard. Then a low complexity but efficient adaptive Golomb-Rice encoder [12] similar to LOCO-I [33] standard is used for encoding the run-length values. For DC coefficients, the redundancy between the consecutive blocks is reduced by using DPCM coder before transforming them into run-length levels. The inverse transform is expressed by:-

$$\hat{X} = C_i\left(X_{Quantized} \circ \hat{\ } \right) \tag{4-13}$$

Where C_i is the core transform of the inverse transformation and $\hat{Q}_i$ is the quantization matrix incorporated with the scaling matrix S_i for the inverse transform. It can be expressed as:-

$$\hat{Q}_i = \left(S_F \circ \ \circ \hat{\ } \right) \tag{4-14}$$

where

$$C_i = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} \text{ and } S_i = \begin{bmatrix} \frac{1}{4} & \frac{1}{2\sqrt{10}} & \frac{1}{4} & \frac{1}{2\sqrt{10}} \\ \frac{1}{2\sqrt{10}} & \frac{1}{10} & \frac{1}{2\sqrt{10}} & \frac{1}{10} \\ \frac{1}{4} & \frac{1}{2\sqrt{10}} & \frac{1}{4} & \frac{1}{2\sqrt{10}} \\ \frac{1}{2\sqrt{10}} & \frac{1}{10} & \frac{1}{2\sqrt{10}} & \frac{1}{10} \end{bmatrix} \quad (4\text{-}15)$$

4.2.4 Corner Clipping

The human gastrointestinal tract is cylindrical in shape, while the lens in the capsule has a circular shape. However, the sensor array captures the image in a rectangular grid. As a result, there are black areas in the corner of the image, which does not contain any diagnostic information. To discard these pixels from processing, the corner clipping mechanism proposed in [17] is utilized. The mechanism is modified to use in block-based coding. With the 2×2 YEFD colour transformation followed by 4×4 linear transformation of individual colour components, the proposed algorithm effectively works on 8×8 blocks. For each 8×8 block, the corner clipping algorithm [17] first calculates the number of pixels in the non-corner region in that block. After the calculation, blocks with a number of pixels in the non-corner region greater than zero are processed, while blocks with zero non-corner pixels are dropped. Figure 4-6 demonstrates the corner clipping procedure.

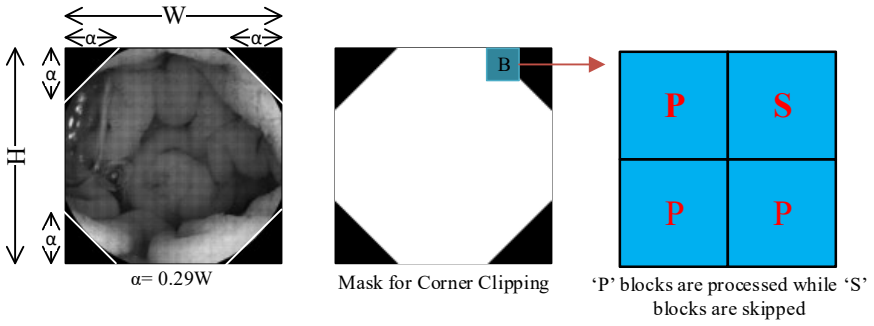


Figure 4-6. Modification of the corner clipping mechanism for block based coding

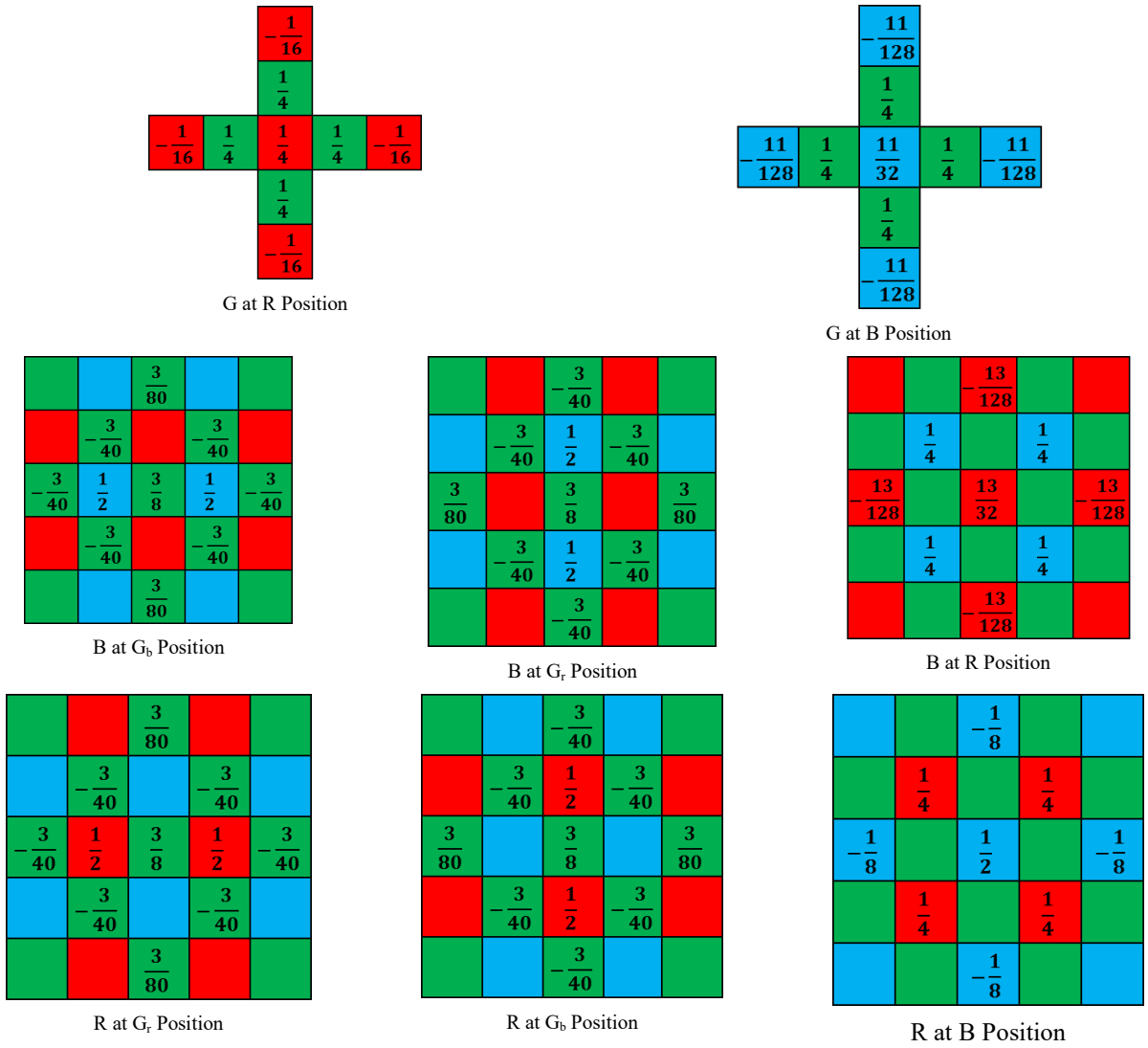


Figure 4-7: Proposed demosaicking filters: (Top) Filter coefficients to interpolate missing green components, (Middle) filter coefficients to interpolate red components, (Bottom) filter coefficients to interpolate blue components

4.2.5 Demosaicking

As stated previously, the compression-first scheme allows the use of high complexity demosaicking algorithm that can effectively reconstruct the colour image as well as reduce the artifacts produced by the compression algorithms. This section proposes a modified high-quality linear interpolation for capsule endoscopy (HQLI-CE) filter to capture the unique characteristics

of the capsule endoscope image and suppress the artifacts resulted by the compression algorithm. The filter followed the similar principle of high- quality linear interpolation (HQLI) filter proposed in [80]. First, using bilinear interpolation, an approximation of the missing colour components are generated. Then a correction term generated from the gradient term are added with the approximation terms to produce the final colour values. Both the bilinear interpolation and gradient correction can be achieved using the linear filtering of the raw CFA pixels. The optimum filter coefficients are calculated using a dataset of capsule endoscopy image [61] using the Weiner approach. The resultant filter coefficients are shown in Figure 4-7. Unlike the original HQLI derived for demosaicking the natural image [80], the filter coefficients differ from the R components and B components.

4.3 Experimental Results

Experiments are carried out using 1943 RGB images from the KID database [9]. This database contains a wide variety of full resolution CE image from different locations with varying conditions. All images were taken using Mirocam and stored using a near-lossless compression method. Therefore, they are suitable for assessment of the proposed algorithm. The full-colour RGB images in the database are sampled by the Bayer CFA to produce the simulated CFA grayscale image. The CFA images are then processed by the proposed pipeline and compressed into bitstream format. The reconstructed CFA images are generated by applying the decoding algorithm to the compressed data. The RGB image is reconstructed by applying the proposed demosaicking algorithm.

In order to evaluate the quality of image reconstruction, the Peak Signal to Noise Ratio is used. The colour transformation was assessed based on local coding gain and entropy. For assessing the performance of the demosaicking algorithm, colour peak signal to noise ratio (CPSNR) and structural similarity index (SSIM) are used.

4.3.1 Colour Transformation

This section compares the performance of the proposed YEFD colour transformation to the original colour channels as well as two colour transformations used in [12]. They will be denoted in this chapter as YYCoCg and YYCuCv colour transformations to distinguish them from their full-colour counterpart YCoCg and YCuCv colour transformations.

Table 4-3: Comparison of coding gain and local coding gain of proposed colour space with other colour space used in compression system for Endoscopic Image

Colour Transform	Coding Gain, Db	Local Coding Gain			
		2x2 Block	4x4 Block	8x8 Block	16x16 Block
Image dependent KLT	**6.79**	1.68	2.32	3.02	3.91
YYCoCg	3.88	1.12	1.31	1.41	1.55
YYCuCv	4.28	1.17	1.42	1.62	1.86
Original Eqn.	6.03	1.98	2.72	3.41	3.80
Proposed Colourspace	6.27	**2.31**	**2.86**	**3.41**	**4.05**

First, the coding gain (G_{TC}), as defined in eqn (4-8), is computed. It represents the reduction in quantization noise by quantizing the transformed domain. To capture the improvement in block-based coding, the local coding gain (LCG) is also computed by computing the variance from NxN non-overlapping blocks and using the mean of the standard deviation in the formula defined in eqn. (4-3).

Table 4-3 shows that the proposed colour space has a significantly higher coding gain than the other colour transformation. This observation holds true for the block based local coding gain. It implies that the condensation of the variance into one luminance channel Y reduces the variance in the chrominance channels L, M, and N significantly in the proposed colour transformation. Therefore, it is expected that more efficient compression can be achieved by compressing the low variance chrominance channels.

Table 4-4: Comparison of Energy Packing Efficiency of the color space transformation

	RGB	YYCuCv	YYCoCG	Proposed
EPE_1	17.70	44.85	45.96	**86.04**
EPE_2	72.79	89.55	91.80	**98.11**
EPE_3	82.38	90.98	93.10	99.84

Table 4-4 demonstrates that the proposed color transformation effectively packs 86% of the energy into one luminance channel. As both YYCuCv and YYCoCg contains two luminance channel, their energy packing efficiency is lower than the proposed color transformation.

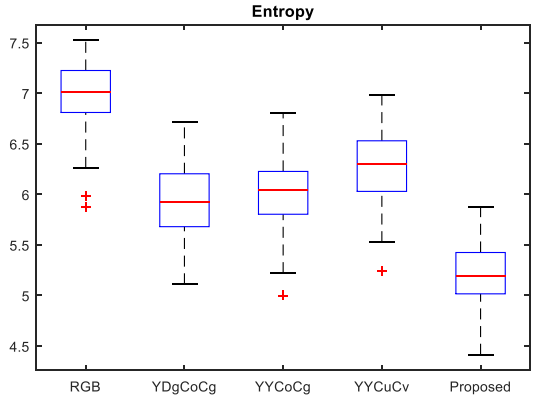


Figure 4-8: Boxplot of the Entropy in the dataset for different color space

Figure 4-8 shows the boxplot for the entropy of the color components for 100 images taken from KID database. Entropy is a measure of information in the data and is the theoretical minimum bound for lossless coding. The entropy of an image can be determined using the equation:-

$$H = -\sum_{i=1}^{n} P_i \log_2 P_i \tag{4-17}$$

here, P_i is the probability of the occurrence of intensity level i. The entropy is expressed in terms of bits per pixel (bpp). We have measured the entropy of each sub-images and taken the weighted mean of the entropy to get the entropy of each image. The weight was measured based on the size of the sub-image compared to the original CFA image. For example, for the proposed color space, where each sub-image is a quarter of the original image, the weight is ¼ for each. While for Green channel in RGB color space and luminance channel in YYCoCg and YYCuCv, the weight is ½. Within each box of the boxplots, the central mark signifies the median and the edge shows the 25th and 75th percentile. Whiskers of each box indicate the 5th percentile and 95th percentile. Outliers are marked as red plus sign. As evident from the boxplot, the proposed color space consistently provides the least entropy. The median of the proposed method is less than the 95th percentile of all the other color space. The proposed color space can yield an average entropy of 5.25 bits per pixel (bpp) while the average entropy of the original color space is 7 bpp.

Figure 4-9. Test Images for Comparing Demosaicking Algorithm

4.3.2 Demosaic Filter Performance

After decoding the CFA image, it is necessary to interpolate the missing colour to generate the full-colour image. In order to investigate the effect of demosaicking technique on the image quality, several interpolation techniques is examined including low complexity techniques such as bilinear interpolation (BI), high quality linear interpolation (HQLI) [80], fully pipelined colour demosaicking (FPCD) [81], and high complexity techniques such as residual interpolation (RI) and adaptive residual interpolation (ARI) [82]. For all these techniques, the original RGB images are first sampled using Bayer pattern and then compressed using the proposed compression system. After decoding, the interpolation techniques were used for regenerating the full-colour image, which is compared with the original RGB image in terms of structural similarity (SSIM) index [83]. SSIM is used to compare the perceptual quality of the image which matches closely with subjective evaluation. Table 4-5 lists the SSIM for 6 test images given in Figure 4-9. To measure the computational complexity, the computational time of the proposed algorithm is compared with fast computational algorithm BI and FPCD. All these algorithms were implemented in MATLAB R2016a and the computational times are obtained on an Intel Core i7 3.4GHz computer. The results are shown in Table 4-6.

Table 4-5: Comparison of Different demosaicking algorithm in terms of SSIM

Image	BI	HQLI	RI	ARI	FPCD	Proposed
1	0.9673	0.9918	**0.9922**	0.9919	0.9881	**0.9926**
2	0.9838	0.9943	**0.9958**	**0.9957**	0.9922	0.9948
3	0.9896	0.9960	**0.9974**	**0.9973**	0.9942	0.9962
4	0.9895	0.9973	**0.9988**	**0.9987**	0.9958	0.9976
5	0.9574	**0.9804**	0.9572	0.9568	0.9503	**0.9815**
6	0.9739	0.9914	**0.9922**	0.9920	0.9874	**0.9921**

As can be seen, the computational time for proposed algorithm is very low and comparable to the BI. The observation shows that the proposed low-complexity demosaicking algorithm can offer a good trade-off between the image quality and computational cost. With a computational complexity slightly higher than bilinear interpolation, this method can provide a comparable subjective image quality to high-complexity algorithms. Therefore, it is suitable to implement in the data recorder for real-time viewing option in WCE system.

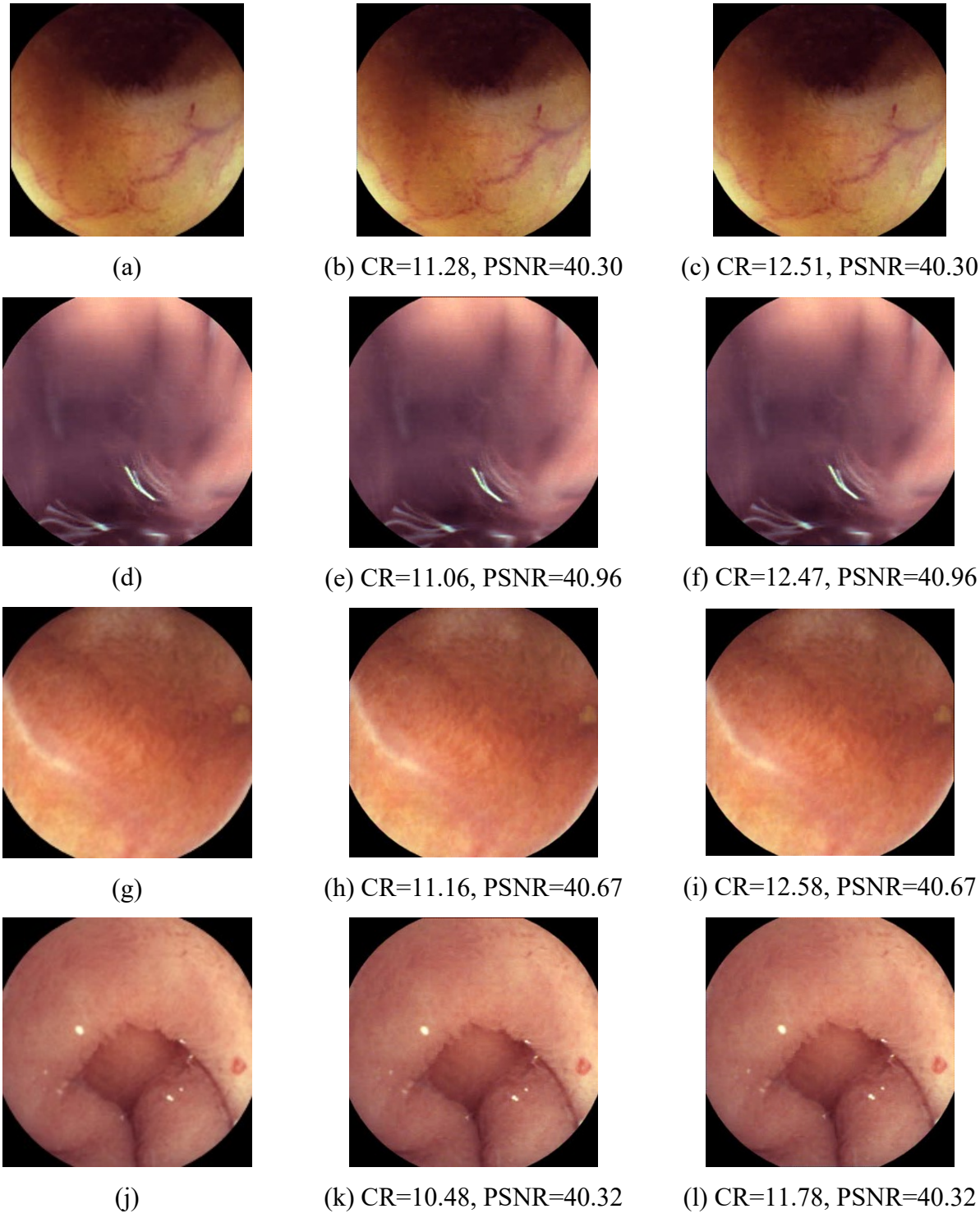
Table 4-6: Comparison of the complexity of low-complexity demosaicking algorithm

	BI	FPCD	Proposed
PSNR	31.29	34.70	**35.96**
SSIM	0.9818	0.9864	**0.9926**
Executing Time	0.034	0.035	0.035

4.3.3 Compression Results

This section compares the coding performance of the proposed compression scheme with JPEG and JPEG-2000 standard engine as well as CFA compression scheme described in the literature. As JPEG and JPEG-2000 operate on the full-colour image only, the RGB image space is interpolated prior to compression. To investigate the effect of low complexity demosaicking algorithm, the bilinear interpolation (BI) algorithm is applied for interpolation. The output compression rate, CPSNR, and SSIM for the KID dataset is shown in Table 4-7. The results clearly illustrate that the standard engine cannot achieve high compression ratio while preserving high image quality due to the use of a sub-optimal demosaicking algorithm. Among the two standard engine, JPEG is superior to JPEG-2000 in terms of compression performance and achievable CPSNR. However, as it can be seen, a higher compression ratio can be achieved by removing the interpolation stage before compression.

The proposed method provides a compression ratio 10.56 on average. Use of the corner clipping increases the average compression ratio rose to 11.83. The image quality is sufficiently high with CPSNR around 40dB and SSIM greater than 0.99. Table 4-8 lists the compression performance reported by other CFA image compression system for capsule endoscopic system. In general, pixel by pixel prediction scheme (e.g. Chen et. al [25]) outperforms the other methods in terms of PSNR and buffer memory requirement.



(a) (b) CR=11.28, PSNR=40.30 (c) CR=12.51, PSNR=40.30

(d) (e) CR=11.06, PSNR=40.96 (f) CR=12.47, PSNR=40.96

(g) (h) CR=11.16, PSNR=40.67 (i) CR=12.58, PSNR=40.67

(j) (k) CR=10.48, PSNR=40.32 (l) CR=11.78, PSNR=40.32

Figure 4-10. Visual comparison of reconstructed images

Methods that works directly on the original colour space [26], [28], [69] offer lower compression efficiency since these methods only reduce the intra-spectral redundancy in the image. As evident from the results in [12], the use of colour transformation significantly increases the compression ratio. However, as it can be seen, higher image quality can be achieved by exploiting the intra-spectral redundancy in CE image by using a dedicated colour transformation such as YEFD. The use of clipping increases the compression ratio closer to 12. A visual comparison of the reconstructed image with the original image is shown in Figure 4-10.

Table 4-7: Compression Results for KID Dataset2

Image Dataset (No of Images)	JPEG			JPEG-2000			Proposed Method			
	CR	CPSNR	SSIM	CR	CPSNR	SSIM	CR without Clipping	CR with Clipping	CPSNR	SSIM
Ampula (19)	4.49	39.1	0.9902	3.25	37.6	0.9945	8.53	9.41	39.3	0.9939
Inflammatory(227)	3.56	39.6	0.9941	2.64	37.8	0.9946	9.52	10.65	39.8	0.9942
Normal Colon (169)	5.47	39.2	0.9879	3.12	38.1	0.9919	9.09	10.09	39.4	0.9925
Normal Esophagus(282)	2.63	39.2	0.9924	2.94	36.8	0.9875	11.69	13.16	40.8	0.9889
Normal Small-bowel(300)	3.53	39.6	0.9939	2.68	37.8	0.9930	12.95	14.75	41.1	0.9900
Normal Stomach (599)	5.87	40.0	0.9913	5.59	39.4	0.9925	10.45	11.63	40.0	0.9941
Polypoid (44)	6.03	39.2	0.9840	3.12	38.1	0.9932	9.51	10.59	39.8	0.9937
Vascular (303)	4.65	39.8	0.9933	2.54	37.6	0.9924	9.25	10.29	39.6	0.9931
Average (1943)	4.47	39.8	0.9918	3.64	38.1	0.9921	10.56	11.83	40.1	0.9924

In addition to the compression performance, Table VIII compares the computational complexity in terms of normalized operation such as addition (a), multiplication (m), shift (s) and division (d). The complexity in terms of the number of operations per pixel before encoding. The comparison illustrates only the method based on algebraic integer quantization (AIQ) based DCT [26] has a lower complexity than the proposed method. The increase in complexity is due to the use of YEFD colour transformation. However, the increase is only 0.3 addition per pixel. Such a marginal increase in computational cost is considered to be tolerable since the use of YEFD colour transformation significantly improves the compression performance.

Table 4-8: Comparison with other CFA compression techniques

Algorithm	Colour Space	CR	PSNR	Computational Complexity	
				Before Encoding	Encoding
Lin et al. [69]	GRBG	4.9	32.5	33a+1s	LZ77
Wahid et al. [26]	GRBG	7.8	32.9	5.5a+1s	LZ77
Dung et al. [28]	GRBG	5.5	36.2	9.4a+4.1s	CAVLC
Chen et al. [25]	GRBG	2.7	46.4	6a+1s	Golomb
Turcza et al. [12]	YCoCgY	11.4	35.7	6a+3s	Golomb
Proposed Without Clipping	**YEFD**	**10.6**	**40.1**	5.8a+3.5s	Golomb
Proposed With Clipping	**YEFD**	**11.8**	**40.1**	5.8a+3.5s	Golomb

4.4 Conclusion

This chapter presents a high-quality lossy compression algorithm to compress Bayer CFA images in WCE system. The proposed system employs a novel colour space transformation to decorrelate the colour components in the CFA image, separates the colour components into four sub-images and compress the sub-images using an integer DCT transform. A novel demosaicking algorithm was proposed to reduce the noise while improving the image quality in the encoder after decompression of the image. Experimental result demonstrates that the proposed scheme effectively reduce both spatial and spectral redundancies, delivering superior compression efficiency in terms of compression ratio and perceptual image quality.

Chapter 5 - Conclusion and Future Work

The advancement of wireless capsule endoscopy would require solving the bottleneck between the image quality and power consumption, as both these plays a major role in the inclusion of active locomotion and therapeutic capabilities. The intention of my research was to contribute to this research area by designing techniques to optimize the performance of both lossy and lossless low complexity image compression algorithm for the next generation wireless capsule endoscopy system.

In Chapter 2, a novel optimization method was introduced for the derivation of reversible lossless colour transformation that works on CFA image. The reversible colour transformation reduces the inter-colour correlation in the prediction error signal to provide error signal with lower variance and lower entropy. The optimization method was used for the derivation of two separate reversible colour transformation of natural image: ORCT1 and ORCT2. These have varying level of computational complexity and compression performance. ORCT1 works directly on raster order data therefore can be implemented without any buffer memory. On the other hand, ORCT2 yields the optimum compression performance with a negligible increase in the computational complexity. The results show that proposed method can provide an optimum colour transformation with a very low computational complexity. It is also demonstrated that the proposed method can improve the prediction based coding by reducing the inter-colour correlation and the entropy of the signal. Experiments with standard lossless compression engine such as JPEG-LS, JPEG-2000 and JPEG-XR demonstrated that the proposed colour transformations provide the superior performance compared to other colour transformation for CFA image. The optimization method is used in Chapter 3 to develop optimized colour transformation for CE image.

In Chapter 3, the goal was to design an optimized lossless compression algorithm for WCE. A low complexity DPCM prediction model was used along with an optimized reversible colour transformation YLMN to exploit the redundancy in both spatial and spectral domain from the CE image in CFA format. The result is a low variance error signal with low inter-colour

correlations. The error signal was then entropy encoded using a single context adaptive Golomb-Rice encoder. A lossless corner clipper mechanism was used to discard the corner regions in the CE image that has no diagnostically relevant information. The results showed that the proposed compression model has a very low computational complexity as it avoids the computationally expensive demosaicking algorithm. Experiments with previous works demonstrated that the use of optimized colour transformation enables the proposed method to achieve the best lossless compression performance with the lowest implementation cost.

In Chapter 4, a novel on cost-effective lossy CFA compression algorithm was proposed for WCE. Each of the stage in lossy compression pipeline right from the colour transformation to demosaicking in the decoder was optimized for obtaining the best compression result for WCE. First, a low complexity colour transformation of different CFA components was derived in CE image using the principal component analysis (PCA). Each of the colour components is then transformed using a 4x4 integer DCT transform and quantized using optimized quantization table. The quantization table was optimized for low complexity and highest performance, where all the divisions are achieved using shift operation. In order to remove the dark corner regions from processing and transmission in the CE image, a modified corner clipping mechanism for the block based coding was employed. Finally in the decoder, a high-quality linear filter was used. The coefficient of the filter was optimized for demosaicking the full-colour image from lossy CFA image and was utilized for high fidelity reconstruction of the full-colour image. Experimental results showed that the proposed colour and linear transformation has an excellent energy packing efficiency. It also demonstrates that the quantized table can selectively omit the uninformative frequencies and provide a very high image compression. Comparison with previous state-of-the-art compression algorithms shows that the proposed method can outperform all of them in terms of compression rate and image quality.

5.1 Contribution

The contribution of this thesis is summarized as follows:-

- A novel reversible colour derivation model is developed for CFA image. The main advantage of the model is that it incorporates the prediction model in the derivation and hence can be used for deriving the optimized colour transformation for a specific prediction model in a specific application. This model reduces the inter-colour correlation in the CFA image

thereby enabling the independent encoding of colour components in order to achieve optimal compression performance. In this way, it removes the need for high computational inter-colour correlation reducing mechanism proposed in the lecture. As the resultant colour transformation is lossless, it can be easily integrated with any prediction based lossless compression algorithm to improve the compression rate.

- Two new reversible colour transformations for lossless encoding of natural images have been proposed. These colour transformations have very low computational complexity and excellent prediction gain performance. Integration of these colour transformations with the standard lossless engine has shown significant improvement over other colour transformations.
- A new reversible colour transformation YLMN was derived for lossless encoding of capsule endoscopic image. An image compression pipeline based on YLMN was developed, which yield the best lossless compression performance both in terms of computational complexity and compression performance for CE image.
- A lossy compression pipeline optimized for coding the CE image in Bayer CFA format has been developed. An optimized colour transformation YEFD has been derived by taking the principal component analysis (PCA) of the different colour components of the CFA images taken from a large dataset of endoscopic images. The optimized quantization table for transformed coefficient from each colour components has been developed. The quantization table has a very low computational complexity as all the division operations have been replaced with the shift operations. A modified high-quality linear filter demosaicking algorithm has been derived to reconstruct the colour image from the lossy CFA data in the decoder.

5.2 Future Work

Looking to the future, the lossless system can be incorporated with a lossy compression system to improve the power consumption as well as the diagnostic yield of WCE. To further improve the image compression system the following research directions are proposed:

- The proposed reversible colour transformation derivation model that works offline when no analytical solution is available for the optimization. An analytical solution would allow the online optimization of the reversible colour

transformation, which can be useful in adaptively changing the colour transformation for different locations in the gastrointestinal tract. One may try to find a simplified optimization model with an analytical solution that can be incorporated in the image compression system to find the optimized colour transformation adaptively.

- The lossless compression method utilizes the DPCM prediction model which has the ability to work in raster order. One may also try to find the optimum prediction model by taking the prediction model parameter as a parameter in the optimization model.
- In this thesis, the lossy and lossless compression algorithms are designed separately. These methods share some common blocks such as colour transformation, DPCM, adaptive Golomb-Rice and corner clipping. A progressive lossy to lossless system, similar to the proposed for JPEG-XR [48], can be adapted to WCE.
- In this thesis, the work was concerned with reducing the redundancy of a single frame. More compression can be achieved by exploiting the redundancy between consecutive frames. One may try to incorporate colourization based video coding as proposed in [84] and [85] along with the proposed lossy compression algorithm.
- Finally, for accurate assessment of the image compressor, image compression algorithm can be implemented in CE prototype and tested in a real-world situation.

References

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

[2] A.-M. Singeap, C. Stanciu, and A. Trifan, "Capsule endoscopy: The road ahead," *World J. Gastroenterol.*, vol. 22, no. 1, pp. 369–378, 2016.

[3] "PillCam SB3." [Online]. Available: http://www.givenimaging.com/en-int/Innovative-Solutions/Capsule-Endoscopy/Pillcam-SB/PillCam-SB-3/Pages/default.aspx. [Accessed: 01-Jan-2017].

[4] "ENDOCAPSULE 10 System-Olympus America –Medical 2016." [Online]. Available: http://medical.olympusamerica.com/products/endocapsule-10-system. [Accessed: 01-Jan-2016].

[5] "OMOM Capsule Endoscopy System." [Online]. Available: http://jinshangroup.gmc.globalmarket.com/products/details/omom-capsule-endoscopy-system-4543846.html. [Accessed: 01-Jan-2016].

[6] "MiroCam." [Online]. Available: http://www.intromedic.com/eng/sub_products_2.html. [Accessed: 01-Jan-2016].

[7] "CapsoVision." [Online]. Available: http://www.capsovision.com/. [Accessed: 01-Jan-2016].

[8] J. A. Barkin and J. S. Barkin, "Video Capsule Endoscopy: Technology, Reading, and Troubleshooting," *Gastrointest. Endosc. Clin. N. Am.*, vol. 27, no. 1, pp. 15–27, 2017.

[9] D. K. Koulaouzidis, A. & Iakovidis, "KID, a capsule endoscopy database for medical decision support," 2016. [Online]. Available: http://is-innovation.eu/kid.

[10] G. Ciuti, A. Menciassi, and P. Dario, "Capsule endoscopy: from current achievements to open challenges.," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 59–72, 2011.

[11] S.-L. Chen, Y.-R. Chen, T.-L. Lin, and Z.-Y. Liu, "A cost-efficient lossless compression color filter array images VLSI design for wireless capsule endoscopy," *J. Med. Imaging Heal. Informatics*, vol. 5, no. 2, pp. 378–384, 2015.

[12] P. Turcza and M. Duplaga, “Hardware-Efficient Low-Power Image Processing System for Wireless Capsule Endoscopy,” *Biomed. Heal. Informatics, IEEE J.*, vol. 17, no. 6, pp. 1046–1056, 2013.

[13] G. K. Wallace, “The JPEG still picture compression standard,” *Commun. ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[14] M. Weinberger, G. Seroussi, and G. Sapiro, “The {LOCO-I} lossless image compression algorithm: Principles and standandization into {JPEG-LS},” *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1309–1326, 2000.

[15] A. Skodras, C. Christopoulos, and T. Ebrahimi, “The JPEG 2000 Still Image,” *IEEE Signal Process. Mag.*, vol. 18, no. September, pp. 36–58, 2001.

[16] F. Dufaux and G. Sullivan, “The JPEG XR Image Coding Standard,” *IEEE Signal Process. Mag.*, vol. 26, no. MMSPL-ARTICLE-2009-004, pp. 195–200, 2009.

[17] T. Khan and K. Wahid, “Lossless and Low-Power Image Compressor for Wireless Capsule Endoscopy,” *VLSI Des.*, vol. 2011, pp. 1–12, 2011.

[18] A. Mostafa, K. Wahid, and S.-B. Ko, “An efficient YUV-based image compression algorithm for wireless capsule endoscopy,” *2011 24th Can. Conf. Electr. Comput. Eng.*, pp. 000943–000946, 2011.

[19] K. A. Fante, B. Bhaumik, and S. Chatterjee, “Design and Implementation of Computationally Efficient Image Compressor for Wireless Capsule Endoscopy,” *Circuits, Syst. Signal Process.*, 2015.

[20] G. Liu, G. Yan, S. Zhao, and S. Kuang, “A complexity-efficient and one-pass image compression algorithm for wireless capsule endoscopy,” *Technol. Heal. Care*, vol. 23, no. s2, pp. S239–S247, 2015.

[21] Y. Gu, X. Xie, G. Li, T. Sun, and Z. Wang, “Two-stage wireless capsule image compression with low complexity and high quality,” *Electron. Lett.*, vol. 48, no. 25, pp. 1588–1589, 2012.

[22] S. Chen, Y. Chen, T. Lin, and Z. Liu, “A Cost-Efficient Lossless Compression Color Filter Array Images VLSI Design for Wireless Capsule Endoscopy,” vol. 5, no. 2, pp. 378–384, 2015.

[23] X. Xie, L. Guolin, X. Chen, B. Ch, and S. Had, “Digital wireless endoscopy capsule system,” pp. 5–8, 2004.

[24] X. Xie, G. Li, X. Chen, X. Li, and Z. Wang, “A Low-Power Digital IC Design Inside the Wireless Endoscopic Capsule,” *IEEE J. Solid-State Circuits*, vol. 41, no. 11, pp. 2390–2400, 2006.

[25] X. Chen, X. Zhang, L. Zhang, N. Qi, H. Jiang, and Z. Wang, “A wireless capsule endoscopic system with a low-power controlling and processing ASIC,” *IEEE Trans. Biomed. Circuits Syst.*, vol. 3, no. 1, pp. 11–22, 2009.

[26] K. Wahid, S. Ko, and D. Teng, “Efficient Hardware Implementation of an Image Compressor for Wireless Capsule Endoscopy Applications,” in *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on. IEEE*, 2008, pp. 2761–2765.

[27] M.-C. Lin, L.-R. Dung, and P.-K. Weng, “An ultra-low-power image compressor for capsule endoscope.,” *Biomed. Eng. Online*, vol. 5, p. 14, Jan. 2006.

[28] L.-R. Dung, Y.-Y. Wu, H.-C. Lai, and P.-K. Weng, “A modified H.264 intra-frame video encoder for capsule endoscope,” *2008 IEEE Biomed. Circuits Syst. Conf.*, pp. 61–64, 2008.

[29] X. Xie, G. Li, X. Li, Z. Wang, C. Zhang, D. Li, and L. Zhang, “A new approach for near-lossless and lossless image compression with Bayer color filter arrays,” *Multi-Agent Secur. Surviv. 2004 IEEE First Symp.*, pp. 357–360, 2004.

[30] X. Xie, G. Li, X. Chen, X. Li, and Z. Wang, “A Low-Power Digital IC Design Inside the Wireless Endoscopic Capsule,” *IEEE J. Solid-State Circuits*, vol. 41, no. 11, pp. 2390–2400, 2006.

[31] X. I. E. Xiang, L. I. Guolin, and W. Zhihua, “A Low Complexity and High Efficient Method for Image Compression with Bayer CFAs *,” vol. 12, no. 1, pp. 22–29, 2007.

[32] S.-L. Chen, Y.-R. Chen, T.-L. Lin, and Z.-Y. Liu, “A cost-efficient lossless compression color filter array images VLSI design for wireless capsule endoscopy,” *J. Med. Imaging Heal. Informatics*, vol. 5, no. 2, pp. 378–384, 2015.

[33] M. J. Weinberger, G. Seroussi, and G. Sapiro, “The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS,” *Image Processing, IEEE Transactions on*, vol. 9, no. 8. pp. 1309–1324, 2000.

[34] X. Xie, G. Li, X. Chen, X. W. Li, B. Ch, S. Han, JinKeYao, C. Zhang, and Z. Wang, “A novel low power IC design for bi-directional digital wireless endoscopy capsule system,”

in *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, 2004, p. S1/8-S5-8.

[35] X. H. Van, J. Ascenso, and F. Pereira, "HEVC backward compatible scalability: A low encoding complexity distributed video coding based approach," *Signal Process. Image Commun.*, pp. 1–20, 2015.

[36] C. Koh, J. Mukherjee, and S. Mitra, "New efficient methods of image compression in digital cameras with color filter array," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1448–1456, 2003.

[37] M. Lin, "A Cardinal Image Compressor for Capsule Endoscope," pp. 146–149, 2006.

[38] P. Turcza and M. Duplaga, "Low-Power Image Compression for Wireless Capsule Endoscopy," in *Imaging Systems and Techniques, 2007. IST '07. IEEE International Workshop on*, 2007, pp. 1–4.

[39] P. Turcza and M. Duplaga, "Low power FPGA-based image processing core for wireless capsule endoscopy," *Sensors Actuators, A Phys.*, vol. 172, no. 2, pp. 552–560, 2011.

[40] S. Lee and A. Ortega, "A novel approach of image compression in digital cameras with a Bayer color filter array," in *IEEE International Conference on Image Processing. Proceedings*, 2001, pp. 482–485.

[41] N. Lian, L. Chang, V. Zagorodnov, and Y. Tan, "Reversing demosaicking and compression in color filter array image processing: performance analysis and modeling.," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3261–78, 2006.

[42] D. K. Iakovidis and A. Koulaouzidis, "Software for enhanced video capsule endoscopy: challenges for essential progress," *Nat. Rev. Gastroenterol. Hepatol.*, vol. 12, no. 3, pp. 172–186, 2015.

[43] G. Sharma and H. J. Trussell, "Digital color imaging," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 901–932, 1997.

[44] R. Lukac and K. N. Plataniotis, "Single-sensor camera image compression," *IEEE Trans. Consum. Electron.*, vol. 52, no. 2, pp. 299–307, 2006.

[45] N. Zhang and X. Wu, "Lossless compression of color mosaic images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1379–1388, 2006.

[46] S. Kim and N. Cho, "Lossless Compression of Color Filter Array Images By Hierarchical," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 1, pp. 1040–1046,

2014.

[47] K. Chung and Y. Chan, “A Lossless Compression Scheme for Bayer Color Filter Array Images,” *IEEE Trans. Image Process.*, vol. 17, no. 2, pp. 134–144, 2008.

[48] H. Malvar and G. Sullivan, “Progressive-to-lossless compression of color-filter-array images using macropixel spectral-spatial transformation,” in *Data Compression Conference Proceedings*, 2012, pp. 3–12.

[49] N. Jayant and P. Noll, *Digital Coding of Waveforms Principles and Application to Speech and Video*. 1984.

[50] H. Malvar, G. Sullivan, and S. Srinivasan, “Lifting-based reversible color transformations for image compression,” in *Proc. of SPIE, Applications of Digital Image Processing XXXI*, 2008, vol. 7073, pp. 707307-707307–10.

[51] M. Adams, F. Kossentini, and R. Ward, “Generalized S transform,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2831–2842, 2002.

[52] T. Khan and K. Wahid, “Low Power and Low Complexity Compressor for Video Capsule Endoscopy,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 10, pp. 1534–1546, 2011.

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

[54] A. Karargyris and N. Bourbakis, “Wireless capsule endoscopy and endoscopic imaging: a survey on various methodologies presented.,” *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 1, pp. 72–83, 2010.

[55] A. Menciassi and G. Ciuti, “Future Developments of Video Capsule Endoscopy: Hardware,” in *Video Capsule Endoscopy*, Springer Berlin Heidelberg, 2014, pp. 543–556.

[56] B. E. Bayer, “Color Imaging Array,” 1976.

[57] D. Hudesman, J. Mazurek, and A. Swaminath, “Capsule endoscopy in Crohn’s disease: Are we seeing any better?,” *World J. Gastroenterol.*, vol. 20, no. 36, pp. 13044–13051, 2014.

[58] C. Spada, M. E. Riccioni, R. Urgesi, and G. Costamagna, “Capsule endoscopy in celiac disease,” *World J. Gastroenterol.*, vol. 14, no. 26, pp. 4146–4151, 2008.

[59] T. H. Khan and K. A. Wahid, “Design of a Lossless Image Compression System for Video Capsule Endoscopy and Its Performance in In-Vivo Trials,” *Sensors*, vol. 14, no. 11, pp.

20779–20799, 2014.

[60] M. Hernandez-Cabronero, I. Blanes, A. J. Pinho, M. W. Marcellin, and J. Serra-Sagrista, "Progressive lossy-to-lossless compression of DNA microarray images," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 698–702, 2016.

[61] "Gastrolab," 2016. [Online]. Available: http://www.gastrolab.net/. [Accessed: 01-Jan-2016].

[62] D. Lee and K. N. Plataniotis, "Lossless compression of HDR color filter array image for the digital camera pipeline," *Signal Process. Image Commun.*, vol. 27, no. 6, pp. 637–649, 2012.

[63] S. Kim and N. I. Cho, "Hierarchical prediction and context adaptive coding for lossless color image compression," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 445–449, 2014.

[64] K. A. Fante, B. Bhaumik, and S. Chatterjee, "Design and Implementation of Computationally Efficient Image Compressor for Wireless Capsule Endoscopy," *Circuits, Syst. Signal Process.*, vol. 35, no. 5, pp. 1–27, 2015.

[65] S. W. Golomb, "Run-Length Encodings," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 1689–1699, 1965.

[66] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *IEEE Trans. Consum. Electron.*, vol. 46, no. 4, pp. 1103–1127, 2000.

[67] J. Li and Y. Deng, "Fast Compression Algorithms for Capsule Endoscope Images," *2009 2nd Int. Congr. Image Signal Process.*, pp. 1–4, 2009.

[68] P. Turcza and M. Duplaga, "Low-Power Image Compression for Wireless Capsule Endoscopy," *Imaging Syst. Tech. 2007. IST '07. IEEE Int. Work.*, pp. 1–4, 2007.

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

[70] M.-C. Lin and L.-R. Dung, "A Subsample-Based Low-Power Image Compressor for Capsule Gastrointestinal Endoscopy," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, p. 257095, 2011.

[71] R. Koprowski, "Overview of technical solutions and assessment of clinical usefulness of capsule endoscopy.," *Biomed. Eng. Online*, vol. 14, no. 1, p. 111, Jan. 2015.

[72] D. Turgis and R. Puers, “Image compression in video radio transmission for capsule endoscopy,” *Sensors Actuators A Phys.*, vol. 123–124, no. April, pp. 129–136, 2005.

[73] X. Chen, X. Zhang, L. Zhang, N. Qi, H. Jiang, and Z. Wang, “A wireless capsule endoscopic system with a low-power controlling and processing ASIC,” *Proc. 2008 IEEE Asian Solid-State Circuits Conf. A-SSCC 2008*, vol. 3, no. 1, pp. 321–324, 2008.

[74] T. H. Khan and K. Wahid, “Low-complexity colour-space for capsule endoscopy image compression,” *Electron. Lett.*, vol. 47, no. 22, p. 1217, 2011.

[75] “Kodak Lossless True Color Image Suite.” [Online]. Available: http://r0k.us/graphics/kodak/.

[76] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[77] H. S. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, “Low-complexity transform and quantization in H.264/AVC,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 598–603, 2003.

[78] H. S. Malvar, “Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts,” *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 1043–1053, 1998.

[79] V. Ratnakar and M. Livny, “An efficient algorithm for optimizing DCT quantization,” *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 267–270, 2000.

[80] H. S. Malvar, L. H. L. He, and R. Cutler, “High-quality linear interpolation for demosaicing of Bayer-patterned color images,” *2004 IEEE Int. Conf. Acoust. Speech, Signal Process.*, vol. 3, pp. 5–8, 2004.

[81] S. Chen and H. Chang, “Fully Pipelined Low-Cost and High-Quality Color Demosaicking VLSI Design for Real-Time Video Applications,” *IEEE Trans. Circuits Syst. II Express Br.*, vol. 62, no. 6, pp. 588–592, 2015.

[82] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, “Residual Interpolation for Color Image Demosaicking,” *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1288–1300, 2016.

[83] Z. Wang and A. C. Bovik, “Structural similarity based image quality assessment,” *Digit. Video Image Qual. Percept. Coding*, p. 20, 2005.

[84] T. H. Khan, S. K. Mohammed, M. S. Imtiaz, and K. A. Wahid, “Color Reproduction and Processing Algorithm based on Real-time Mapping for Endoscopic Images,”

Springerplus, pp. 1–10.

[85] T. H. Khan, S. Member, S. K. Mohammed, and S. Member, “Efficient Color Reproduction Algorithm for Endoscopic Images based on Dynamic Color Map,” pp. 16–17.

Appendix A

This thesis is primarily based on several manuscripts resulting from the collaboration of multiple researchers. All publications have been modified to make the thesis coherent.

For the manuscripts in Chapter 2 and Chapter 4, the author's contribution is in developing the idea, implementing the algorithm, designing the experiments and writing the manuscript. Dr. Khan A. Wahid guided through the research and provided helpful suggestions.

The work in Chapter 3, is the result of the collaboration of multiple researchers. The author's contribution is in developing the idea, designing the lossless compression algorithm, corner clipping algorithm, designing the experiments and writing the manuscript. KM Rahman helped in the hardware implementation of the image compression algorithm in ASIC and FPGA. Dr. Khan A. Wahid guided throughout the research and provided helpful suggestions.

Appendix B

List of Publications

B.1 Published Peer Reviewed Journals

1. T. Khan, **S. Mohammed**, M. Imtiaz and K. Wahid, “Efficient Colour Reproduction Algorithm for Endoscopic Images Based on Dynamic Colour Map," Journal of Medical and Biological Engineering, 2016, 36(2), 226-235.
2. T. Khan, **S. Mohammed**, M. Imtiaz and K. Wahid, “Colour Reproduction and Processing Algorithm Based on Real-time Mapping for Endoscopic Images," SpringerPlus, 2016, 5(1), 1-17.
3. R. Shrestha, **S. Mohammed**, M. Hasan, X. Zhang and K. Wahid, “Automated Adaptive Brightness in Wireless Capsule Endoscopy System Using Image Segmentation and Sigmoid Function," IEEE Transactions on Biomedical Circuits and Systems, 2016, 10(4), 884-892.

B.2 Published Conference Papers

1. **Shahed K. Mohammed**, Farah Deeba, Francis M. Bui, and Khan A. Wahid. “Feature Selection Using Modified Ant Colony Optimization for Wireless Capsule Endoscopy.” The 7th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON 2016), New York City, USA.
2. **Shahed K. Mohammed**, Farah Deeba, Francis M. Bui, and Khan A. Wahid. “Application of Modified Ant Colony Optimization for Computer Aided Bleeding Detection System,” in The 2016 International Joint Conference on Neural Networks (IJCNN 2016), pp. 4317–4324, 2016, Vancouver, Canada.
3. Farah Deeba, **Shahed K. Mohammed**, Francis M. Bui, and Khan A. Wahid. “Unsupervised Abnormality Detection Using Saliency and Retinex based Colour Enhancement.” 38th Annual International Conference on IEEE Engineering and Medicine and Biology Society (EMBC’16), pp. 3871–3874, 2016, Orlando, FL.

4. Farah Deeba, **Shahed K. Mohammed**, Francis M. Bui, and Khan A. Wahid. “An Empirical Study on the Effect of Imbalanced Data on Bleeding Detection in Endoscopic Video.”38th Annual International Conference on IEEE Engineering and Medicine and Biology Society (EMBC’16), pp. 2598–2601, 2016, Orlando, FL.
5. Farah Deeba, **Shahed K. Mohammed**, F. M. Bui, and K. A. Wahid, “A saliency-based unsupervised method for angioectasia detection in capsule endoscopic images," 2016 Canadian Medical and Biology Engineering Society Conference (CMBEC39), Calgary, Canada.