2014 IEEE 28-th Convention of Electrical and Electronics Engineers in Israel

Low Complexity Image Compression of Capsule Endoscopy Images

Aviv Barabi, Dvir Sason and Rami Cohen
Signal and Image Processing Lab (SIPL)
Department of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200000, Israel

***Abstract*—Capsule endoscopy is a method for recording images of the digestive tract. A patient swallows a capsule containing a tiny camera, which captures images that are then transmitted wirelessly to an external receiver for examination by a physician. The images are captured using a Bayer filter mosaic, such that each pixel in raw captured images represents only one color: red, green or blue. Due to limited computational capabilities in the capsule and bandwidth constraints, low-complexity and efficient compression of Bayer endoscopy images is required before transmission. In this paper, we focus on the JPEG algorithm, and show how to adapt it for compressing Bayer images. We show that by transformation to the YCgCo color space and appropriate optimization of parameters, significant improvement is achieved over the standard JPEG.**

I. Introduction

Capsule endoscopy [1] is a state-of-the-art method for recording images of the digestive tract. This medical procedure is performed by swallowing a pill-size capsule containing a tiny camera, which captures and transmits images of the gastrointestinal tract. Capsule endoscopy is especially efficient for detecting polyps, ulcers and tumors of the small intestine. With the advance of this technology, the images are transmitted in real time to an external receiver, to be reviewed by a physician.

However, the transmission rate of video content during the endoscopy process is limited due to limited power resources and bandwidth constraints. The small dimensions of the capsule limit its computational capabilities, resulting in restriction on resolution and quality of the transmitted images. Usually, the images are captured using a Bayer filter mosaic [2] where only partial color information is available. Traditional image compression techniques, such as JPEG [3], operate on full color images, and have poor performance when applied to raw Bayer images. Therefore, efficient compressing schemes that meet the capsule constraints and are adapted to the Bayer format are desired.

Several methods were suggested for compressing Bayer endoscopy images. In [4], it was proposed to combine the integer versions of the discrete cosine transform (DCT) and the discrete wavelet transform (DWT) for improved efficiency and reduced complexity. Compressed sensing based algorithm was proposed in [5], where the YUV color space is considered. Reduction of encoding complexity was considered in [6], by distributing the complexity between the encoder and the decoder using statistical models. However, these methods require the development of new compression schemes whose complexity might be difficult to analyze, and which are not necessarily modular.

In this work, we focus on the existing JPEG standard, and show how it can be modified for compressing Bayer endoscopy images for providing good compression performance under complexity constraints. Due to its good performance and relatively low complexity, The JPEG format is one of the most common image formats used by digital cameras, allowing a selectable tradeoff between compressed image size and quality.

By representing the raw Bayer images in the RGB and YCgCo color spaces, and using designated run-length encoding (RLE) Huffman tables, we show how a significant improvement in performance can be achieved. Due to complexity considerations, we propose to use integer DCT (ICT) instead of the DCT, where we show that no degradation in performance is incurred. Later, we formulate an optimization problem, for the determination of the optimal quantization factors under a quality constraint.

The paper is structured as follows. In Section II, we discuss several modifications of JPEG for the compression of Bayer endoscopy images. An optimization problem for the determination of optimal quantization factors is formulated in Section III. Experimental results are provided in Section IV. Finally, conclusions are given in Section V.

II. Adaptation of JPEG to Bayer Endoscopy Images

The images captured by the capsule are acquired using a Bayer filter mosaic, such that each pixel in the captured images represents only one color: red, green or blue. Bayer images are composed of 2×2 blocks, where the two diagonal pixels are green and the remaining two are red and blue. To obtain full color images, a demosaicing algorithm [7] is applied to the raw Bayer image. In this work, we used the demosaicing algorithm suggested in [8], which is based on low complexity interpolation and provides good results.

In principle, one may compress the full color images after the demosaicing process, due to higher correlation between pixels in the full color image. However, this approach is obviously inefficient, since it requires demosaicing at the

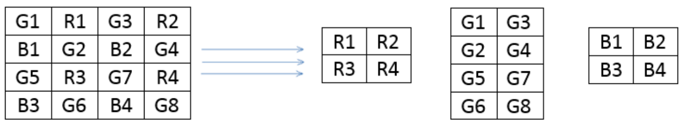


Fig. 1: RGB decomposition of a Bayer format image.

capsule and compression of a larger amount of information. Thus, we consider here compression of raw Bayer endoscopic images.

Different demosaicing algorithms result in different reconstructed color images. In addition, several post-processing methods may be applied to the reconstructed images, for improved image display. Therefore, we will measure the reconstruction quality in the Bayer domain. The quality measure we will use is the peak signal-to-noise ratio (PSNR) measured in dB, defined as:

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{255^2}{\mathrm{MSE}}, \tag{1}$$

where MSE denotes the mean squared-error between the original Bayer endoscopy image and its reconstructed version. Note that 8-bit images are assumed.

Due to complexity constraints, we will not exploit temporal information (i.e., correlation between neighbouring frames in a capsule endoscopy video), and each image will be treated separately. Therefore, we will not use here techniques such as motion estimation, which is the major tool used in video compression. Instead, we will rely on spatial correlation only, as is done in the JPEG standard. In the following, we discuss several compression tools, which will be later combined for the compression of Bayer endoscopy images.

A. RGB Color Space

Since neighbouring pixels in Bayer images represent different colors, their correlation is low compared to full color images. The aim of the DCT used in the JPEG algorithm is to exploit spatial correlation between pixels, so that compressing the raw Bayer image is ineffective. For that reason, one of the preprocessing methods we considered is the decomposition of the Bayer images into 3 images, each representing one of the colors: red, green or blue.

The 2 green value pixels in a block are arranged diagonally, so we shifted the lower green value pixel in each 2×2 block to obtain a vertical arrangement of the green pixels. An example is shown in Fig. 1. Finally, each color component was compressed separately. This way, the spatial correlation of each color channel is exploited.

B. YCgCo Color Space

Capsule endoscopy images are usually characterized by a dominating red color, since they are taken in inner body parts. Moreover, Bayer images are composed of 2 times more green value pixels than red and blue, where the green color contributes most to luminance information. This unique color distribution pattern is well captured by representing the Bayer image in the YCgCo color space [4].

To transform a Bayer image into YCgCo color space we used the following transformation as suggested in [4]:

$$\begin{pmatrix} Y^{ul} \\ Y^{lr} \\ Cg \\ Co \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/4 & 1/4 \\ 0 & 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 0 & 0 & -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} G^{ul} \\ G^{lr} \\ B \\ R \end{pmatrix}, \tag{2}$$

where G^{ul} is the upper left green pixel in each 2×2 block, G^{lr} is the lower-right green pixel, B is the blue pixel and R is the red pixel. Accordingly, Y^{ul} is the upper-left luma component, Y^{lr} is the lower-right luma component, C_g stands for the green chroma pixel and C_o stands for the orange chroma pixel. Note that the transformation matrix coefficients in Eqn. (2) are all powers of 2, enabling an efficient implementation in hardware. Similarly to the compression of the RGB components, we evaluated the performance of JPEG encoding when applied to each color separately.

C. Integer DCT

The DCT is used in JPEG, especially because of its strong energy compaction. In the H.264/MPEG-4 AVC standard [9], a more hardware efficient version of DCT called integer DCT (ICT), is used. This transform requires multiplication of the pixel values by powers of two, which can be implemented efficiently using shift operations. Varying block sizes such as 4×4, 8×8 and 16×16 are used in the H.264/MPEG-4 AVC standard, instead of 8×8 blocks in the JPEG standard.

As described in [10], the ICT merges the quantization process with the transform process. Therefore, the use of ICT in JPEG replaces both the transform and quantization stages. Note that when applying ICT, higher quantization factors correspond to rougher quantization. We concentrate here on 4×4 and 8×8 block sizes, since 16×16 blocks are too large for the typical low resolution of capsule endoscopy images.

D. Pre-Learnt Huffman tables

In the JPEG baseline algorithm, each quantized non-zero AC coefficient is represented by two run-length encoding (RLE) symbols. The first symbol contains the number of non-zero AC coefficients preceding it and the number of bits required for encoding it. The second symbol contains the amplitude of the AC coefficient. DC coefficients are encoded in a similar manner, but since they are differentially encoded, their corresponding first symbol contains amplitude only.

For improved compression, the RLE symbols are encoded using Huffman coding. There are standard Huffman tables that provide the mapping between each possible RLE symbol and its corresponding Huffman codeword. These tables were generated by averaging the Huffman tables of many images. However, these images were full color images, which do not represent correctly the structure of Bayer images. Therefore, improvement in performance is expected when Huffman tables adapted for the Bayer format are used.

We compared two approaches for obtaining Bayer Huffman tables. In the first one, which we name "online learning", the Huffman table is generated and transmitted for each image. This way, we get the optimal (in the sense of close to entropy compression performance) Huffman table for each image. In the second approach, which we name "offline learning", the Huffman tables are pre-learnt using a set (training set) of Bayer images. This approach does not require the transmission of Huffman tables with each image, since they are known at the decoder side.

In both approaches, a histogram of each RLE symbol (for the DC and the AC coefficients) is generated for each image. The frequency of each symbol in the histogram is then stored, to be used later for Huffman coding. In the online learning approach, the Huffman tables of each image are transmitted as the header of the compressed image. On the other hand, in the offline learning approach, the statistics of the training set images were merged for the calculation of Huffman tables that represent the average statistics of the training set. As we will see later, the offline learning approach is sufficient in our case, since the statistics of Bayer capsule endoscopy images are quite similar.

III. Optimization Problem Formulation

As mentioned in Section II-B, different colors in endoscopic images are not equally important. Therefore, different color channels require different treatment, i.e., different quantization factors, according to the importance of the colors to the reconstructed image quality. To get the appropriate quantization factor for each color component, expressions for PSNR and bit-per-pixel (bpp) as a function of the quantization factor of each color component are desired. Such expressions are difficult to find, so we will concentrate here on approximations.

In the following, the YCgCo color space is considered, but the RGB color space can be used instead. For simplicity, we shifted Y^{lr} to the left, such that the Y channel combines Y^{ul} and Y^{lr} (similarly to the G component in Fig. 1). The bpp of the total Bayer image is:

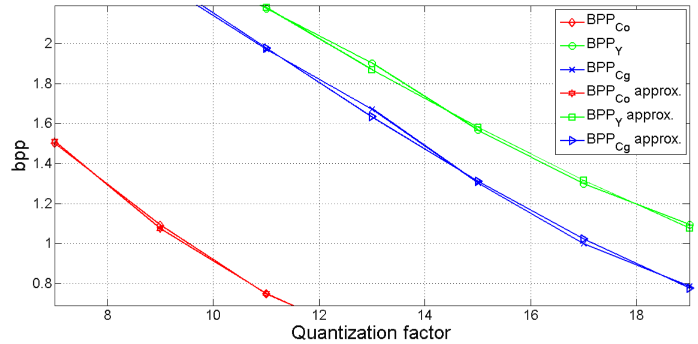
$$\mathrm{bpp} = \frac{1}{2}\mathrm{bpp_Y} + \frac{1}{4}\mathrm{bpp_{Cg}} + \frac{1}{4}\mathrm{bpp_{Co}}, \qquad (3)$$

where $\mathrm{bpp_Y}$ is the bpp of the Y channel, $\mathrm{bpp_{Cg}}$ is the bpp of the Cg channel and $\mathrm{bpp_{Co}}$ is the bpp of the Co channel. The coefficients are the proportions of the Y, Cg or Co pixels out of the total number of pixels. In a similar manner, the overall PSNR as a function of the PSNR of each color channel is:

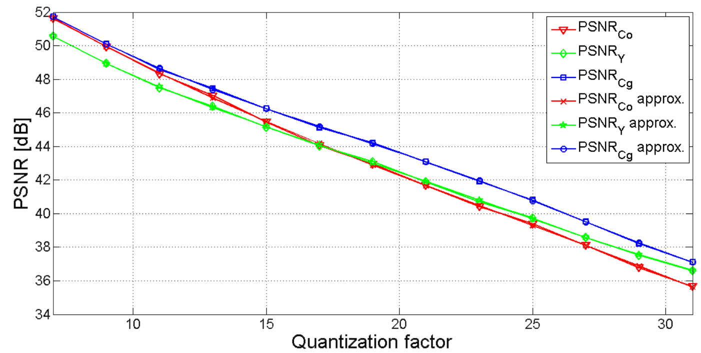
$$\mathrm{PSNR} = \qquad (4)$$
$$-10 \cdot \log\left(\frac{1}{2}\cdot 10^{-\frac{\mathrm{PSNR_Y}}{10}} + \frac{1}{4}\cdot 10^{-\frac{\mathrm{PSNR_{Cg}}}{10}} + \frac{1}{4}\cdot 10^{-\frac{\mathrm{PSNR_{Co}}}{10}}\right),$$

where $\mathrm{PSNR_Y}$ is the PSNR of the Y channel, etc.

The function we would like to minimize is the bpp as defined in Eqn. 3, where the quality constraint is a PSNR value higher than a predefined threshold. The variables are three quantization factors: Q_Y, Q_{Cg} and Q_{Co} that should be determined for each color component separately. The resulting optimization problem is:



(a) bpp.



(b) PSNR.

Fig. 2: bpp and PSNR approximation as degree 5 polynomials.

$$\begin{array}{ll} \text{Variables} & Q_Y, Q_{Cg}, Q_{Co} \\ \text{minimize} & \mathrm{bpp}\,(Q_Y, Q_{Cg}, Q_{Co}) \\ \text{s.t.} & \mathrm{PSNR}\,(Q_Y, Q_{Cg}, Q_{Co}) \geq \text{Threshold}. \end{array} \qquad (5)$$

For solving this optimization problem, we approximated the empirical bpp and PSNR curves as a function of each quantization factor to degree 5 polynomials, using sample Bayer images. This approximation turned out to be very effective, as can be seen in Fig. 2. These polynomials were then used for providing expressions for bpp and PSNR in the optimization problem, using Eqns. (3) and (4).

The optimization problem in (5) can now be solved analytically using Lagrange multipliers or optimization tools such as CVX [11]. It is expected that different color components will be assigned different quantization factors, due to the different proportions of the color components in the Bayer image. In case of non-integer optimal quantization factors, rounding should be used. Note that for improved performance, the bpp and PSNR polynomials can be updated throughout the transmission, using information from already reconstructed images.

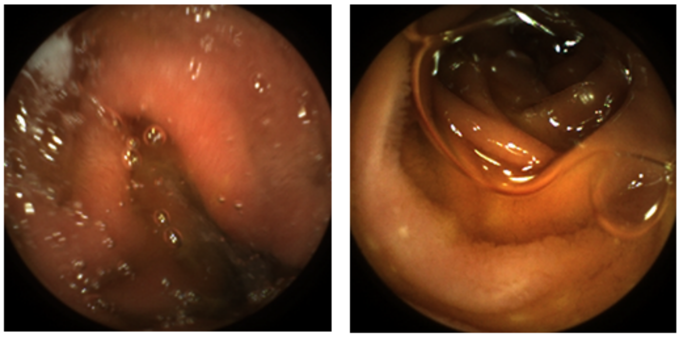


Fig. 3: Two sample capsule endoscopy images (after demosaicing).

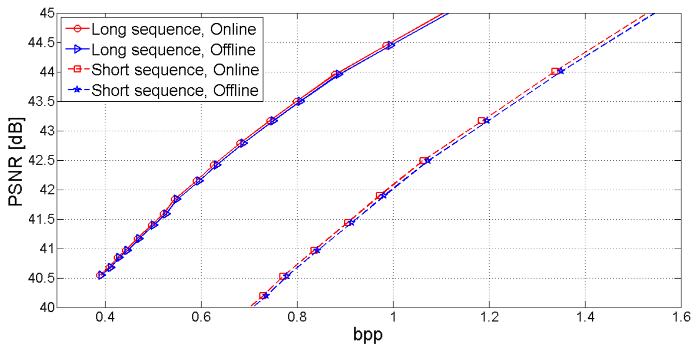


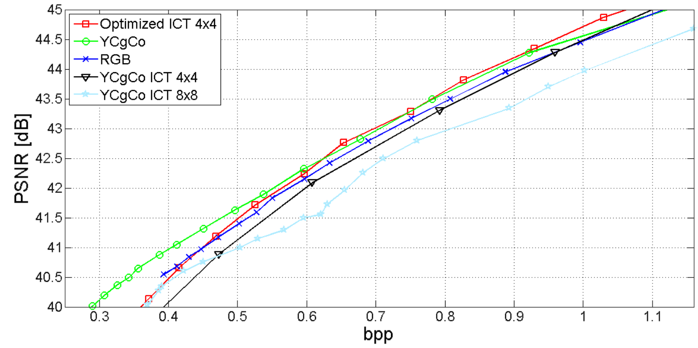
Fig. 4: Comparison of online/offline learning of JPEG Huffman tables.

IV. Experimental Results

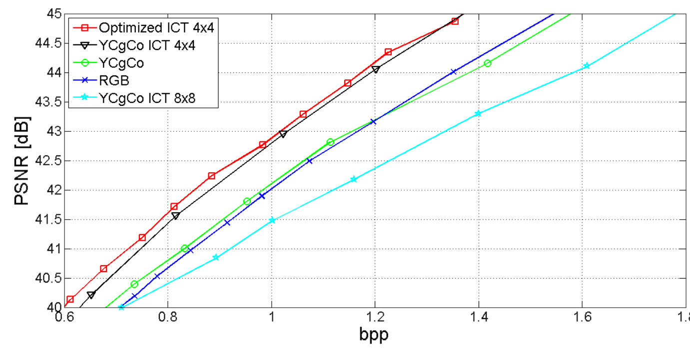
We evaluated the performance of the suggested compression methods using two sequences of Bayer capsule endoscopy images with resolution of 320×320. The first sequence ("long sequence") is composed of 346 images, while the other one ("short sequence") is composed of 140 images. Both sequences show parts of the small intestine, as shown in Fig. 3.

As noted in Section II-D, we compared the "online/offline learning" approaches for learning Huffman tables. As an example, a comparison between the learning schemes is presented in Fig. 4, where the RGB decomposition was considered. We can see that for both videos the PSNR gain is about only 0.1dB when online learning is used. Due to the additional complexity required in the online learning, we preferred to use offline learning for the remaining methods.

The results obtained using the methods discussed in Section II and the optimization problem of Section III are shown in Fig. 5. These results show that the use of 4×4 ICT with optimized quantization factors for the YCgCo color space provides the best performance. This maintains the low complexity requirement, since the transformation to the YCgCo color space and the ICT can be implemented by simple shift operations.



(a) Long sequence.



(b) Short sequence.

Fig. 5: Experimental results.

V. Conclusion

In this paper, we provided several modifications of the JPEG algorithm for the compression of Bayer endoscopic images. By learning the JPEG Huffman tables for Bayer endoscopy images, an improvement of 3dB in performance was obtained when a straightforward decomposition of Bayer images to RGB components was considered. We saw that offline learning of the Huffman tables is sufficient, meaning that it can be performed outside the capsule.

An additional improvement in performance of 0.5–1dB was achieved by a transformation to the YCgCo color space and an optimization of the quantization factors. This optimization requires the solution of a simple optimization problem, which can be obtained offline. The compression methods considered in this paper are based on simple arithmetic operations, and can be implemented efficiently in hardware.

Acknowledgement

The authors would like to thank the staff of the Signal and Image Processing Lab (SIPL) for their support.

References

[1] D. O. Faigel and D. R. Cave. *Capsule Endoscopy*. Elsevier, 2008.

[2] B. E. Bayer. Color imaging array. *U.S. Patent 3,971,065*, 1976.

[3] G. K. Wallace. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 38(1):xviii–xxxiv, Feb 1992.

[4] P. Turcza, T. Zielinski, and M. Duplaga. Low complexity image coding algorithm for capsule endoscopy with bayer color filter array. In *Signal Processing Algorithms, Architectures, Arrangements and Applications, 2007*, pages 27–32, Sept 2007.

[5] J. Wu and Y. Li. Low-complexity video compression for capsule endoscope based on compressed sensing theory. In *IEEE Engineering in Medicine and Biology Society (EMBC) Conference*, pages 3727–3730, Sept 2009.

[6] R. Cohen and D. Malah. LORD: LOw-complexity, Rate-controlled, Distributed video coding system. In *Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS)*, pages 21–28, Nov 2012.

[7] R. A. Maschal Jr. et al. Review of Bayer Pattern Color Filter Array (CFA) Demosaicing with New Quality Assessment Algorithms. Technical report, U.S. Army Research Laboratory, 2010.

[8] H. S. Malvar, L.-W. He, and R. Cutler. High-quality linear interpolation for demosaicing of bayer-patterned color images. In *International Conference of Acoustic, Speech and Signal Processing (ICASSP)*, May 2004.

[9] D. Marpe, T. Wiegand, and G. J. Sullivan. The H.264/MPEG4 advanced video coding standard and its applications. *IEEE Communications Magazine*, 44(8):134–143, Aug 2006.

[10] I. Richardson. 4x4 transform and quantization in H.264/AVC, 2010.

[11] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1, 2014.

[12] K.-H. Chung and Y.-H. Chan. A lossless compression scheme for bayer color filter array images. *IEEE Transactions on Image Processing*, 17(2):134–144, Feb 2008.

[13] C. Ramirez, F. J. Mora Mas, and J. Enseat. Property of images in bayer formats in the digital processing of images. In *Electronics, Robotics and Automotive Mechanics Conference (CERMA)*, pages 267–271, Sept 2008.

[14] S.-Y. Lee and A. Ortega. A novel approach of image compression in digital cameras with a bayer color filter array. In *International Conference on Image Processing*, volume 3, pages 482–485, Oct 2001.