Colour-reproduction algorithm for transmitting variable video frames and its application to capsule endoscopy

Tareq Khan, Ravi Shrestha, Md. Shamin Imtiaz, Khan A. Wahid ✉

Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, Saskatchewan S7N5A9, Canada
✉ E-mail: khan.wahid@usask.ca

Published in Healthcare Technology Letters; Received on 16th September 2014; Revised on 5th January 2015; Accepted on 7th January 2015

Presented is a new power-efficient colour generation algorithm for wireless capsule endoscopy (WCE) application. In WCE, transmitting colour image data from the human intestine through radio frequency (RF) consumes a huge amount of power. The conventional way is to transmit all R , G and B components of all frames. Using the proposed dictionary-based colour generation scheme, instead of sending all R , G and B frames, first one colour frame is sent followed by a series of grey-scale frames. At the receiver end, the colour information is extracted from the colour frame and then added to colourise the grey-scale frames. After a certain number of grey-scale frames, another colour frame is sent followed by the same number of grey-scale frames. This process is repeated until the end of the video sequence to maintain the colour similarity. As a result, over 50% of RF transmission power can be saved using the proposed scheme, which will eventually lead to a battery life extension of the capsule by 4–7 h. The reproduced colour images have been evaluated both statistically and subjectively by professional gastroenterologists. The algorithm is finally implemented using a WCE prototype and the performance is validated using an ex-vivo trial.

1. Introduction: Wireless capsule endoscopy (WCE) presents an appealing alternative to traditional wired endoscopy for gastrointestinal (GI) diagnosis, preventing discomfort to the patient. However, it has limited working time, lower frame rate and lower image resolution because of its power limitation compared with its counterpart. Therefore one key design challenge of WCE development is to transmit and receive high-quality video frames using as little power as possible. Moreover, the next-generation WCE system could deploy multi-band imaging and therapeutic robotic features that will perform biopsy, microsurgery, drug delivery and tissue sampling [1, 2]. All of these additions will require more battery power of the electronic capsule.

In conventional WCE, colour images are captured by the electronic capsule and then transmitted out through a radio frequency (RF) waves to the data logger where the images are stored for diagnosis. Normally, commercial image sensors convert the Bayer images into RGB images and transmit them out [3]. Transmitting all three colour components (i.e. red, green and blue) of a colour image consumes significant processing time and battery power. To save transmission power, several image compression algorithms have been proposed [2, 4, 5] that reduce the number of bits needed to be transmitted from the capsule. Although the image compression algorithm helps the cause, it itself consumes a small amount of power and hardware resources. In this Letter, a novel way is presented that can save a significant amount of transmission power without the need for additional hardware resources. Here, a new dictionary-based algorithm for generating colour images is proposed that reproduces colour images from a grey-scale image (or luminance component). As a result, instead of transmitting all colour images from the electronic capsule, it will only transmit one colour image followed by a series of grey-scale images. The number of grey-scale images used can be controlled by the user or the system (typically, three or four images for every colour image). The proposed algorithm is suitable for image compressors that work on luma-chroma colour spaces (such as YUV, YCbCr etc. [2, 4]). It is adaptive by nature and will be implemented in the data recorder where more hardware and power resources are available. Since no additional hardware is required in the electronic capsule and grey-scale image sequence is mostly transmitted, a

significant amount of RF power can be saved, which results in an extended battery life of the capsule. Fig. 1 shows the illustration of the proposed method against the conventional method.

The motivation of the proposed adaptive frame transmission comes from the fact that the GI tract has homogeneous colour on its mucosa surface within a short distance. Moreover, inside the small intestine, the capsule travels by peristalsis action at a slow speed of about 1–2 cm/min or 0.16–0.32 mm/s and sends images at 2–3 frames/s [1]. Thus, the capsule sends 6–12 frames within a travel distance of 1 mm. Owing to colour homogeneity, these frames will very likely have similar colour tone. As a result, if we transmit one colour image at the beginning followed by three (or four or five) grey-scale images and later colourise the grey-scale images with the tone of the first colour image, there will be no colour distortion in the colourised images, or if there is, it will be insignificant. This proposed technique has been verified statistically and experimentally during ex-vivo trial, and later evaluated by professional gastroenterologists. In all cases, high similarity between the original colour image and the colourised images has been achieved.

2. Related works on colour generation: Colour reproduction refers to assigning a 3D colour vector to a 1D luminance or brightness realisation. There are many colour generation

Figure 1 Illustration of proposed frame transmission
Shown for case 1:4, i.e. one colour frame followed by four grey or luma frames

algorithms found in the literature that mainly add colour tone from a colour image to a grey-scale image [6]. In [7], the colour information is assigned to each pixel in the grey image by matching each grey-scale image pixel to a pixel in the target swatch using the Euclidean distance matrix. In another work, [8], the authors have used a set of seed points and their respective colour vectors in the *RGB* format with a YUV-based classification. In [9], a quadratic objective function-based optimisation method is used to interpolate the U and V components of the YUV colour space over the entire image using a set of colour scribble lines. In [10], pseudo colours are employed to colourise the grey-scale image using different 64×3 colour matrices. None of the above mentioned methods have been applied to endoscopic imaging. In a recent work by the present authors' group, a colour enhancement scheme is presented that is dedicated to enhancing endoscopic images [11].

In this Letter, we propose an adaptive colour generation method that can reproduce the original colour tone from a theme image without introducing additional colour artefacts. First of all, a colour map dictionary is dynamically generated from a theme colour image as each input colour pixel is read sequentially. The table is indexed by the luminance (or luma) values present in the image. Then, the new colour information is added from the dictionary to the matched luma pixels on a grey-scale image to produce a colourised image. The colour map is updated every time a new colour image is read and the process continues. The algorithm has much lower computational complexity and execution time which makes it very suitable for low-power applications.

3. Proposed colour generation method: The proposed colour generation scheme has three steps. The steps are briefly discussed below.

3.1. Colour map generation: A colour map is a function, Z , implemented as a look-up table, which maps luminance to colour components as the following equation

$$Z:Y \rightarrow \{R, G, B\} \quad (1)$$

Here, the colour map is generated from a colour image, referred to as 'theme' image, T . The reproduction stage generates an output image, O , having the similar colour of the image T . To generate the colour map from T , each colour pixel of the image is converted to luminance, Y , using (2) and its corresponding R , G and B values are recorded in a look-up table or dictionary indexed by the Y value (all ranges from 0 to 255)

$$Y = \frac{R}{4} + \frac{G}{2} + \frac{B}{4} \quad (2)$$

The dictionary contents are dynamic and adaptive as its contents change whenever the theme image is changed. A sample look-up table is shown in Fig. 2. While filling in the dictionary entries, an

Y	R	G	B
0	0	0	0
1	3	0	0
2	2	0	2
3	6	1	2
4	4	3	4
⋮			
255	255	255	255

Figure 2 Generation of colour map dictionary

exception may occur where different combinations of R , G and B would produce the same Y value. For example, *RGB* triplets $\{16, 8, 4\}$ and $\{0, 14, 0\}$ will both produce a Y value of 7, according to (2). In this case, the R , G , B values that are already in the look-up table are averaged with the new colour values using (3) where X represents R , G and B ; N is the total number of combinations of R , G and B that would produce the same Y value in the theme image

$$X = \frac{1}{N} \sum_{n=1}^N X_n \quad (3)$$

Moreover, after the dictionary is generated, there might be some empty slots for some Y values. These slots are filled by linear interpolation between the two nearest neighbouring entries. For instance, if there are some consecutive empty slots from row M to row $M + d$ in the look-up table, then the empty slot entries are calculated using (4), where X represents R , G , B and i range from 0 to d

$$X_{M+i} = X_{M-1} + \frac{X_{M-1} - X_{M+d+1}}{d+2} \times (i+1) \quad (4)$$

The first and last entries of the table are kept fixed presenting full black $\{0, 0, 0\}$ and full white $\{255, 255, 255\}$ namely. A pseudo code for colour map generation algorithm is given in Fig. 3.

3.2. Smoothing of R , G , B entries using curve fitting: After generating the colour map dictionary, it is found that the R , G and B entries are scattered in random and non-continuous fashion with respect to Y . This may produce artefacts in the colourised image since a small change in Y will produce large colour differences with respect to its neighbouring pixels. To minimise this effect, a linear interpolation-based curve fitting algorithm is implemented in the second stage. Fig. 4a shows the plot of the generated colour map before the smoothing where the first frame from an endoscopic video sequence (shown in Fig. 5) is used as the theme image. Fig. 4b shows the *RGB* plot after smoothing.

3.3. Applying the colour map: In the last stage, the colour map is added to the grey-scale frames, E , to produce the output

```

For Each Pixel in T
{
  Y := Get_Y (R, G, B)
  If TblEntryExist (Y) = False
  { //Fill Table
    Tbl_R (Y) := R;
    Tbl_G (Y) := G;
    Tbl_B (Y) := B;
    TblEntryExist (Y) := True; }
  Else { //Average old and the current value
    Tbl_R (Y) := (Tbl_R (Y) + R)/2 ;
    Tbl_G (Y) := (Tbl_G (Y) + G)/2 ;
    Tbl_B (Y) := (Tbl_B (Y) + B)/2 ; }
}
//Interpolate Empty Slots
For i = 1 To 255
{If TblEntryExist(i) = False
{b_i := i - 1;
  For j = i + 1 To 255
  {If TblEntryExist(j) = True
  {e_i := j; Exit For ;}}
  d := e_i - b_i;
  dR := (Tbl_R(e_i) - Tbl_R(b_i)) / d;
  dG := (Tbl_G(e_i) - Tbl_G(b_i)) / d;
  dB := (Tbl_B(e_i) - Tbl_B(b_i)) / d;
  For n = 1 To d - 1
  {Tbl_R(b_i+n) := Abs(Tbl_R(b_i)+dR*n);
    Tbl_G(b_i+n) := Abs(Tbl_G(b_i)+dG*n);
    Tbl_B(b_i+n) := Abs(Tbl_B(b_i)+dB*n); }
  i := e_i;
}
}

```

Figure 3 Pseudo code of colour map generation algorithm

Figure 4 Plot of generated colour map and colour map after smoothing
a Plot of generated colour map
b Colour map after smoothing

colourised image, O . Here, the R , G and B values are directly mapped from the colour map dictionary, Z , for each luminance pixel of the grey-scale image, E as expressed in the following equation

$$Z: \forall Y \in E \rightarrow \{R, G, B\} = O \quad (5)$$

4. Results and discussion

4.1. Performance assessment using video frames: The colour generation method has been applied on several endoscopic video sequences collected from [12]. Fig. 5 shows the result of five frames. The colour map dictionary was generated from frame #1 of Fig. 5a and then applied to four consecutive grey-scale frames

Figure 5 Frames from endoscopic video sequence, generated colour image using proposed method and original colour images
a Frames from endoscopic video sequence
b Generated colour frames using proposed method where frame #1 of (a) is used as theme image
c Original colour frames – shown for comparison with (b) only

to produce colour frames as shown in Fig. 5b. Fig. 5c shows the original colour video frames from the video for comparison purpose only. Now comparing Fig. 5b with Fig. 5c, we can see that the generated colour frames have a very similar colour tone and texture to the original.

To validate the results statistically, the colour similarity between the original and colourised images is evaluated using several performance metrics such as the CIE94 delta- E colour difference [13], structure similarity index (SSIM), mean SSIM (MSSIM) [14] and structure and hue similarity (SHSIM) [15]. In CIE94, $\Delta E_{94}^* \approx 2.3$ indicates that the colour difference between two images is the lowest. The indices of SSIM, MSSIM and SHSIM close to 1 indicate higher similarities between two colour structures. The results are compared with other colour generation methods and presented in Table 1. Here, six endoscopic video sequences have been used. In each video sequence, 40 consecutive frames are used in generating the results. Finally, for each performance metric, the average of all frames ($6 \times 40 = 240$ frames) is reported in Table 1. We can see that the SSIM, MSSIM and SHSIM indices are higher than others with a colour difference ΔE_{94}^* is close to 2.3. All these indicate that the colourised frames are very close to the original frames.

In addition to generating close to the original colour, the proposed method has extremely low complexity, which results in very fast execution speed. The time required for colour reproduction of different image sizes using various methods are shown in Table 2. The same database (six video sequences, 40 frames each) is used in this experiment and the average of all frames is reported. The simulations were conducted on a workstation having Intel(R) Core(TM) i3-2310M CPU at 2.10 GHz and 6 GB RAM. The proposed algorithm works in two stages and the time breakdown for both stages is also given in the table. It can be seen that the proposed method is much faster than [7, 11]. Although the execution time of [10] is lower than ours, its quality of colour reproduction is much worse than ours, as seen in Table 1.

4.2. Performance validation using ex-vivo trial: The proposed scheme is further validated using an ex-vivo trial with a WCE prototype developed in our lab [16]. The experimental setup is shown in Fig. 6. Several frame transmission schemes (i.e. one

Table 1 Colour similarity assessment

Similarity between two colour images		ΔE_{94}^*	SSIM	MSSIM	SHSIM
original colour image	proposed	2.37	0.9993	0.9392	0.9987
	[11]	3.1	0.9917	0.8705	0.8178
	[7]	1.6	0.9907	0.8084	0.8997
	[10]	4.93	0.8175	0.6850	0.5049

Table 2 Comparison of processing time with other works

Image size	Method	Time for colour map generation, s	Time for colour reproduction, s	Total processing time, s
256×256	proposed	0.530	0.909	1.439
	[11]	—	28.255	28.255
	[7]	—	64.461	64.461
	[10]	—	0.193	0.193
512×512	proposed	1.966	3.809	5.775
	[11]	—	113.314	113.314
	[7]	—	232.146	232.146
	[10]	—	0.471	0.471

Figure 6 Capsule prototype and ex-vivo trial of proposed scheme
a Capsule prototype (size 16 mm × 36 mm)
b Ex-vivo trial of proposed scheme

colour frame followed by a different number of grey-scale frames) have been used during the trial. All original colour frames were also stored for comparison purposes. The results are shown in Fig. 7. Like the simulation results before, the generated coloured frames have very similar colour tone and texture as the original frames. Fig. 8 shows the coloured frames when other methods are used on the grey-scale frames collected during the trial.

4.3. Subjective evaluation: To verify the acceptability of the coloured frames, we have conducted a quick survey among four medical imaging doctors (names are given in the Acknowledgment section). In the survey, three sets of images (with each set consisting of a grey-scale image, the coloured image and original colour image) were shown to the participants, who were later asked to give their opinion score as to whether they agree with the statement – ‘the generated coloured image can be used instead of the original colour image’. The results of the survey are summarised in Table 3. The average mean opinion score (MOS) [17] is 0.83, which can be ranked as ‘agree’.

It is noted that no firm conclusion can be drawn from the preliminary survey. A more rigorous task-based approach is needed with more medical experts and samples, which is beyond the scope of the Letter [18]. Nevertheless, the survey was just conducted to obtain an initial assessment of the proposed method.

4.4. Analysis of power saving: In this section, we show how much power can be saved using the proposed method for WCE application. It is noted that the colour generation scheme is implemented in the data logger and no additional hardware resources are needed in the capsule to implement it. Let us consider a scenario where the capsule sends a frame sequence that consists of one colour frame followed by N number of grey-scale images. We refer to it as [1: N]. Then the average number of bits (S) to transfer one frame is given by the following equation

$$S = \frac{W \times H}{1 + N} \{BPP_C + (N \times BPP_G)\} \times (1 - CR) \quad (6)$$

where W is the image width, H is the image height, CR is the rate of

Figure 7 Images captured during ex-vivo trial (shown for case 1:3)
a Live frames from capsule
b Color generated frames
c Original color frames (for comparison)

Figure 8 Colour generated frames
a Using [11], $\Delta E_{94}^* = 1.66$
b Using [7], $\Delta E_{94}^* = 1.44$
c Using [10], $\Delta E_{94}^* = 4.09$

compression and BPP_C and BPP_G are the bit width in bits/pixel for colour and grey-scale images, respectively. Now, if E_b is the energy consumption per bit of the RF transceiver, the total energy (E_T) required to transmit S bits is $E_T = S \times E_b$.

From it, the power consumption (P) is calculated using (7) where the frame rate f_r is in frame/s

$$P = E_T \times f_r \quad (7)$$

Now we have calculated the power consumption for the conventional (all colour frames) and proposed cases (with frame sequence varying from 1:1 to 1:5) and the results are shown in Fig. 9. Here, $f_r = 2$ frames/s [1], image size is 256×256 with $BPP_C = 24$,

Table 3 Quality assessment survey results

	Doctor 1, S_1	Doctor 2, S_2	Doctor 3, S_3	Doctor 4, S_4	MOS
image #1	2	0	0	1	0.75
image #2	2	1	1	1	1.25
image #3	1	-1	1	1	0.50
average MOS					0.83

Strongly agree = 2, agree = 1, neutral = 0, disagree = -1, strongly disagree = -2; MOS = $\sum\{S_1:S_M\}/M$, $M = 4$

Figure 9 Savings in power consumption for different frame sequences

$BPP_G = 8$ and $CR = 80, 60$ and 0% for lossy compression [4], lossless compression [5] and no compression, respectively. E_b is measured to be 21.05 nJ/bit [19].

From Fig. 9, we can see that for conventional transmission (i.e. 1:0 – meaning all frames are colour), the RF power consumption is the highest for all cases: 13.24 mW for lossy, 26.48 mW for lossless and 66.21 mW for uncompressed images. Whereas, using the proposed scheme, the power can be significantly reduced for all cases. For example, using a frame sequence 1:3 (i.e. one colour frame followed by three grey-scale frames), the consumption is reduced to 6.62 , 13.24 and 33.10 mW for lossy, lossless and uncompressed cases, respectively, which results in a saving of over 50% .

It would be ideal to use a commercial capsule (such as Pillcam SB2, whose power rating is 77 mW [20]) to show the advantage of the proposed transmission scheme; however, it cannot be done since there is no access to Pillcam's internal hardware and so the default transmission sequence cannot be modified. As a result, we used the capsule prototype we designed in our lab [16] to show the performance and validate our claim on power saving.

The power saved using the proposed scheme may be used to increase the capabilities of the capsule, such as frame size, frame rate, robotic features and increase the battery life. To demonstrate the advantage, we have calculated the battery life keeping all existing parameters constant and using a battery capacity of 190 mAh and 3.3 V. The power consumption of other modules of a WCE system (such as image sensors, LED and controller) is taken from [2], which is 10.56 mW. The estimated battery life for different transmission sequences is shown in Table 4. It can be seen that the power saving achieved in RF transmission results in an extension of 4 – 7 h of battery life. The saved power can be used to increase the image resolution. For example, a capsule with 256×256 image resolution and 2 fps has estimated 8 h of battery life. Using the proposed method, we can attain similar battery life with 320×320 image resolution, 2 fps and 1:1 frame transmission scheme. Alternatively the saved power can be used to increase the frame rate (fps). As a result, more GI tract area will be covered by the capsule, which helps to perform a more accurate and complete diagnosis.

Table 4 Avg. RF power and estimated battery life of endoscopy capsule

Method	Frame seq.	Average RF power consumption, mW	Est. battery life, h ^a
conventional (all colour)	1:0	66.21	8.17
proposed	1:1	44.14	11.46
	1:2	36.78	13.24
	1:3	33.10	14.36
	1:4	30.90	15.12
	1:5	29.43	15.68

^aUsing a battery of capacity 190 mAh and 3.3 V

Figure 10 Savings in power consumption for different frame sequences, resolutions and frame rate (using a battery of capacity 190 mAh and 3.3 V)

A graphical presentation of the estimated battery life using different image resolutions and frame rate is shown in Fig. 10. This will help users determine the configuration to achieve a targeted battery life; for example, a battery life of 8 h can be achieved in two ways: (a) by setting a frame size of 320×320 at 2 fps using 1:1 transmission sequence or (b) by setting a frame size of 256×256 at 4 fps using 1:3 transmission sequence.

5. Conclusion: In this Letter, we have presented an adaptive dictionary-based colour generation scheme that is targeted for capsule endoscopy application. The scheme generates a colour map dictionary from a colour frame and then adds colour information to the following grey-scale frames. As a result, it will enable a number of frame transmission sequences in capsule endoscopy, all of which will save a significant amount of RF power. As a result, the battery life of the capsule may also be increased to facilitate a more accurate and complete diagnosis.

6. Acknowledgments: The authors acknowledge Western College of Veterinary Medicine, Univ. of Saskatchewan for providing animal testing facilities. Thanks to Dr. M. Greer of The Hospital for Sick Children, Dr. P. Ellul of Mater Dei Hospital, Dr. R. Fedorak of Univ. of Alberta, and Dr. S. Wiebe of Univ. of Saskatchewan for participating in the survey.

7. Funding and declaration of interests: This work was supported by research funding from the Natural Science and Engineering Research Council of Canada (NSERC) and Grand Challenges Canada (GCC) Star in Global Health. The equipment and lab infrastructure was supported by Canada Foundation for Innovation (CFI). Conflict of interest: none declared.

8 References

- Toennies J.L., ET AL.: 'Swallowable medical devices for diagnosis and surgery: the state of the art', *J. Mech. Eng. Sci.*, 2009, **224**, pp. 1397–1414
- Khan T., Wahid K.: 'White and narrow band image compressor based on a new color space for capsule endoscopy', *Image Commun. Signal Process.*, 2014, **29**, (3), pp. 345–360
- Khan T., Wahid K.: 'Design of a lossless image compression system for video capsule endoscopy and its performance in *in-vivo* trials', *Sens. MDPI*, 2014, **14**, (11), pp. 20779–20799
- Khan T., Wahid K.: 'Low power and low complexity compressor for video capsule endoscopy', *IEEE Trans. Circuits Syst. Video Tech.*, 2011, **21**, (10), pp. 1534–1546
- Khan T., Wahid K.: 'Lossless and low power image compressor for wireless capsule endoscopy', VLSI design (Hindawi Publishing Corporation, 2011), vol. 2011, Article ID 343787
- Yatziv L., Sapiro G.: 'Fast image and video colorization using chrominance blending', *IEEE Trans. Image Process.*, 2006, **15**, pp. 1120–1129
- Welsh T., ET AL.: 'Transferring color to greyscale images', *ACM Trans. Graph.*, 2002, **21**, (3), pp. 277–280
- Horiuchi T., ET AL.: 'Colorization algorithm for grayscale image by propagating seed pixels'. Proc. of ICIP, 2003, pp. 457–460

- [9] Levin A., *ET AL.*: 'Colorization using optimization', *ACM Trans. Graph. (TOG)*, 2004, **23**, pp. 689–694
- [10] Korostyshevskiy V.: 'Grayscale to RGB converter', 2006. Available at: <http://www.mathworks.com/matlabcentral/fileexchange/13312-grayscale-to-rgb-converter>
- [11] Imtiaz M.S., Wahid K.A.: 'A color reproduction method with image enhancement for endoscopic images'. Proc. IEEE Second Middle East Conf. on Bio-Med. Engineering, Doha, Qatar, February 2014
- [12] Faigel D., Cave D.: 'Capsule endoscopy' (Saunders Elsevier, 2008)
- [13] Robertson A.R.: 'Historical development of CIE recommended color difference equation', *Color Res. Appl.*, 1990, **15**, (3), pp. 167–170
- [14] Wang Z., *ET AL.*: 'Image quality assessment: from error visibility to structural similarity', *IEEE Trans. Image Process.*, 2004, **13**, pp. 600–612
- [15] Shi Y., *ET AL.*: 'Structure and hue similarity for color image quality assessment'. ICECT, 2009, pp. 329–333
- [16] Khan T., Shrestha R., Wahid K.: 'A modular and programmable development platform for capsule endoscopy system', *J. Med. Syst.*, 2014, **38**, (6), p. 12
- [17] Istepanian R., *ET AL.*: 'Subjective and objective quality assessment in wireless teleultrasonography imaging'. Proc. Int. Conf. EMBS, 2008
- [18] Pezzoli A., Cannizzaro R., *ET AL.*: 'Interobserver agreement in describing video capsule endoscopy findings: a multicentre prospective study', *Dig. Liver Disease*, 2011, **43**, (2), pp. 126–131
- [19] Zainul C.: 'Nordic radio nRF24L01 + power characterization', 2014. Available at: <http://www.nesl.ee.ucla.edu/fw/zainul/NordicRadio.pdf>
- [20] Pillcam SB2, Given imaging. <http://www.givenimaging.com/en-us/Pages/default.aspx>