Lossless Image Coding Based on Inter-color Prediction for Ultra High Definition Image

Jiho Park[1], Je Woo Kim[1], Jechang Jeong[2], and Byeongho Choi[1]

[1] Multimedia IP Research Center, KETI,
25 Saenari-ro Yatap-dong, Bungdang-gu, Seongnam-si, Gyeonggi-do 463816, Korea
[2] Department of Electronic and Computer Engineering, Hanyang University,
222, Wangsimni-ro, Seongdong-gu, Seoul 133791, Korea
{scottie,jwkim,bhchoi}@keti.re.kr
jjeong@ece.hanyang.ac.kr

Abstract. This paper addresses the lossless image coding for ultra-high definition television system which supports 4K (4096×2160) resolution image with 22.2ch audio. Ultra High Definition Tele-vision system is being developed to satisfy end-user who has a longing for higher resolution, higher quality, and higher fidelity picture and sound. However, the major characteristic of Ultra High Definition Tele-vision system is considerably huge input information must be processed in real-time compare to conventional systems. Therefore high speed data handling for editing and playing, high speed signal interface between devices and real-time source codecs without delay for saving memory space are unavoidable requirements of ultra-high system. This paper focused on lossless image codec for the reason of real-time processing with reasonable coding gain and the proposed algorithms is pixel based algorithms called spatio-color prediction. The proposed algorithm uses inter-color correlation with spatial correlation appropriately and it shows 5.5% coding efficiency improvement compare to JPEG-LS. The simulation was performed using 4K resolution RGB 10bit images and it is claimed that the proposed lossless coding was verified under the developed Ultra High Definition Tele-vision system.

Keywords: lossless coding, ultra high definition, high-fidelity, inter-color.

1 Introduction

Nowadays, many multimedia products are introduced on consumer electronics market and consumers are looking for high speed electronic devices with both high resolution and better quality more and more. To meet the high demand at the market, a lot of products supporting high quality, high speed, high resolution and high fidelity are being developed by broadcasting equipment industry and professional video equipment industry. Analog TV is very rapidly replaced by HDTV which is one of the most familiar products and the end-user expects better performance products. Therefore, the industry needs to develop the next generation product to keep pace with the evolution of environment.

Y.-S. Ho (Ed.): PSIVT 2011, Part II, LNCS 7088, pp. 1–12, 2011.

The most predictable and feasible next generation media service after prevailing HDTV devices at home consumer market would be focused on high resolution and better quality image service to be provided. World cinema industries have already introduced the production system using digital cameras, which can provide digital contents with 4K (4096×2160) resolution exceeding 2K (2048×1080). Also on broadcasting field, many researches regarding commercialization of Ultra HDTV have already been materialized evacuating from the current HDTV (1920×1080). Japan has announced in 2008 its implementation plan of next generation broadcasting system to be exploiting from 2015 and UK has a plan to broadcast 2012 London Olympic Games by Ultra High Definition Television (UHDTV).

Meanwhile, the international standardization organizations such as ISO/IEC JTC1 SC29WG11 MPEG and ITU-T SG16 Q.6 VCEG have started international standardization procedures on over 2K resolution videos under the name of HEVC (High Efficiency Video Coding).

Under such a circumstance, UHDTV system is needed badly on every media environmental fields including the technology development of Ultra High Definition (UHD) contents to edit, store, and play freely for fulfillment of the UHDTV system. But huge input information can hardly be edited, stored, and played freely at real time under the current media technology, though the real-time process of input information is imperative on UHDTV system.

Also lossless coding would be essential to satisfy users who want to have high quality resolution without losses. The most suitable compression technology at present might be high speed JPEG-LS based-technology which has efficient compression effects on information processing. But current prevailing compression technology can hardly store the high capacity input information and it is needed to develop high efficiency compression algorithms in order to have an appropriate and efficient UHDTV system.

The developed system is a comprehensive total system supporting UHD image which is delivered from 4K digital camera and displayed by four 2K 10bit LCD monitors. Furthermore the system architecture for high speed data storage and optical transmission for high speed interface are under development, too.

The specification of input images used on UHDTV system is RGB 4:4:4 10bit depth and it has totally different characteristics of images from those used on conventional applications such as YUV 4:2:0 8bit depth. Thus, verified algorithms for 10bit depth image with reasonable coding efficiency are required to save storage space and also no delay and minimum complexity is required to process huge input information at the same time. Accordingly this paper has been prepared to improve the compression algorithm which has RGB 4:4:4 10bit depth 4K image and to provide modified JPEG-LS [1]-[3] based algorithms.

Following sections would be consisted of the explanation of the proposed intercolor prediction on section 2, the new context modeling for 4K image on section 3, the description of used test images on section 4, the simulation results on section 5 and then the conclusion on the section 6.

2 The Proposed Lossless Coding

In this section, the description of newly developed lossless coding is provided. The main concept of the proposed algorithms is exploiting spatial correlation of UHD image because 4K image has stronger spatial correlation compared to less 2K images. To improve coding efficiency, spatio-color prediction scheme is introduced which is pixel based algorithms such as [3]-[6] and it brought the concept of SICLIC [7] for predictions.

2.1 Inter-color Prediction

In this paper, all images introduced are using RGB color space and each pixel has 10bit depth, and they are coded in order of *G(reen)→B(lue)→R(ed)* line by line. As *G* color plane has the strongest correlation among the other color planes, the *G* color plane has been coded at the first place, and through this prior coding of *G* color plane, the reconstructed *G* color plane can be used as a reference by *B* and *R* color planes. In this proposed inter-color prediction, all the pixels are coded pixel by pixel and, in addition to that, the same prediction scheme as JPEG-LS is applied to the *G* color plane.

If inter-color correlation between current pixels and reference pixels is strong enough, a reference pixel can be used as a predictor of the current pixel and inter-color correlation can be measured by correlation coefficient as shown in (1).

$$S(ICR,SR)=\frac{M\sum_{i=a}^{d}(ICR_i\times SR_i)-\sum_{i=a}^{d}ICR_i\times\sum_{i=a}^{d}SR_i}{\sqrt{\{M\sum_{i=a}^{d}ICR_i^2-(\sum_{i=a}^{d}ICR_i)^2\}\times\{M\sum_{i=a}^{d}SR_i^2-(\sum_{i=a}^{d}SR_i)^2\}}} \quad (1)$$

In equation (1), *M* indicates the number of pixels used *(=4)*, SR_i indicates the value of neighboring pixels, ICR_i means the value of corresponding reference pixels, and *S* represents the strength of inter-color correlation. The corresponding pixels and neighboring pixels are located as shown in Fig. 1.

Fig. 1. Location of corresponding reference pixels and neighboring pixels

I_x on the Fig.1 indicates the current encoding/decoding pixel and the correlation coefficient can be obtained through utilizing corresponding 4 pixels from the reference color plane. Needless to say, the value of obtained correlation coefficient is (0~1) and the value indicates the guideline of correlation ratio among corresponding color planes.

Inter-color prediction is implemented when the value of correlation coefficient exceeds 0.5 as strong correlation exists between them and the reference pixels used in encoder side are reconstructed and used identically in decoder side as well.

All the pixels are applied for the inter-color prediction when they have strong correlation exceeding 0.5 and the compensation of color level displacement shall be followed to adjust color intensity in each color plane. If it is assumed that the relationship between ICR_x and I_x is linear, ICR_x can be approximated to I_x ($=I'_x$) and it can be treated as the simple linear equation like $P_{ix}=I'_x=W \times ICR_x+O$ where W is weight factor and O is offset, and P_{ix} is inter-color predictor of I_x. In case of lossless coding, both encoder and decoder can have the value of ICR_x identically, and the value of O is always zero when the floating point operation is possible. Therefore, if the value of W is obtainable, the color level displacement can be compensated. Meanwhile, the value of W can be easily acquired through the minimization of MES as (2).

$$W(ICR,SR)=\frac{M\times \sum_{i=a}^{d}(ICR_i\times SR_i)-\sum_{i=a}^{d} ICR_i\times \sum_{i=a}^{d} SR_i}{\sqrt{\{M\times \sum_{i=a}^{d} ICR_i^2-(\sum_{i=a}^{d} ICR_i)^2\}}} \tag{2}$$

2.2 Spatio-color Prediction

The spatial correlation is very high on 4K image. Generally speaking, 4K image has more correlation for both horizontally and vertically compared to less 2K image when the same object is taken by both 2K camera and 4K camera respectively. It cannot be guaranteed double correlation increasing when the image resolution is doubled by horizontal and vertical direction, but it is certain that 4K image has stronger spatial correlation than less 2K image. Therefore, only when the inter-color correlation is high enough, better coding efficiency improvement can be expected from inter-color prediction than spatial prediction. Furthermore, there is possibility that spatial correlation may exist, though prediction is performed among inter-color. Consequently, a new prediction scheme is needed to be adopted.

According to the spatio-color prediction written in this paper, the improvement of coding efficiency could be achieved from each following difference scheme, after classifying all pixels in B and R color planes into following three categories.

```
1.  If CC×PF1 exceeds 0.5
2.  Else if CC×PF2 exceeds 0.5
3.  Otherwise
```

The value of PF applied here is a precision factor and it means the accuracy of the correlation to be applied. The value of PF_1 is 0.501 whereas the value of PF_2 is 0.51, which shows the above case 1 is chosen only when the correlation coefficient exceeds 0.998 and above case 2 is chosen, only when the correlation coefficient exceeds 0.98 whereas above case 3 for the balance. As we can realize from the value of the used threshold, inter-color prediction is applied only when relatively high correlation is kept and coding gain cannot be expected by inter-color prediction from the above case 3. So the same prediction scheme as JPEG-LS is highly recommended in order to prevent coding performance degradation.

In case the current pixel is chosen as the above case 1, unlikely to the conventional method using one predictor either inter-color predictor or spatial predictor, the both are needed for application. For derivation of the new predictor P_{new}, the inter-color predictor can be secured from the corresponding reference pixel ICR_x by multiplying weight factor W with the spatial prediction predictor which can be secured from neighboring pixels avoided an edge. A new spatio-color predictor can be obtained through weighted sum as per following (3) according to characteristics of image.

$$P_{new} = (\alpha)P_{ix} + (1-\alpha)P_x \tag{3}$$

The author has defined the value α, which indicates the characteristics of image, from 0.65 to 0.5 for the purpose of this study. An image like zoom-in has fairly high correlation in spatial, whereas less α value brings better performance result. And high α value brings better performance result, in case of spatially complicated images. In fact, better compression ratio can be expected if we manipulate two values; α and PF, appropriately because the value of α affects reciprocally much to the value of PF_1, but there remains the possibility of precision problem of floating point operation and that of serious complexity increasing by continuous threshold. During the simulation, general performance results can be achieved when the value of α keeps 0.5.

If spatial prediction direction of ICR_x is horizontal,
   $P_{ix} = SR_a + W_1 \times (ICR_x - ICR_a)$
Else if spatial prediction direction of ICR_x is vertical,
   $P_{ix} = \mathrm{SR_b} + W_1 \times (ICR_x - ICR_b)$
Else if sptail prediction direction of ICR_x is diagonal
   $P_{ix} = (SR_a+SR_b-SR_c)+ W_1\times(ICR_x-(ICR_a-ICR_b+ICR_c))$
Else
   $P_{ix} = W_2\times ICR_x$

Fig. 2. Derivation process of inter-color prediction for case 2

In case the current pixel is chosen as the above case 2, inter-color correlation is not strong enough like case 1 though weak inter-color prediction scheme is proposed using inter-color correlation as much as possible. Presumption can be made in this proposed method that there is high possibility to have the same directional prediction between the reference color plane and the current color plane. And also residual of neighboring pixels have similar level even though there exists weak inter-color correlation. Therefore P_{ix} can derive from the newly developed algorithm using the prediction direction of reference color plane as per in fig. 2 and final predictor P_{new} can be acquired through (3). For case 2, the recommended value of α is 0.5.

As like case 1, to compensate color level displacement in case 2, the weight factor could be introduced but the coding gain compared to complexity turned out to be negligible because residual of neighboring pixels by spatial prediction are normally small. Accordingly $W_1=1$ is suggested in this paper and W_2 can derive from (2).

In order to implement the proposed method, prediction direction and prediction error value of reference color plane are needed. It has been confirmed through many experiments that the value of prediction error should be remained in a certain area otherwise it would cause performance degradation. Test images used for simulation have 10bit depth therefore 11bit per pixel is needed for residue. However, it is also confirmed in this paper that prediction residue exceeding -15 or 15 does not bring any coding gain, thus the proposed algorithm clipped the absolute value of prediction residual within 0 ~ 15. In case of prediction direction, five directions - skip, horizontal, vertical, diagonal and inter-color prediction are possible to use. Therefore, one byte per pixel was assigned in memory then the prediction residual of reference color plane is stored in LSB 4bits and prediction direction of reference color plane is stored in MSB 4bits.

Moreover, we could find that the reference color plane of *B* matched with *G* color plane, whereas both *G* color plane and *B* color plane can be reference color plane of *R* color plane when *GBR* coded in turn. But indication is needed to decoder on which reference plane has been used for the reference color plane of *R* color. This indication will bring overhead problem which causes performance degradation in term of the output bitstream size. Both the appointment of reference color plane per line by line and per color plane may be considered to minimize overhead problem. The selective reference color plane method can be employed as an alternative, but serious complexity increasing occurs because the output bitstream would be generated after several iterations by the encoder in order to find the best reference color plane, though decoder would know easily from the parsing of the coded bitstream. For the reason, every reference color plane is restricted within *G* color plane only in order to minimize encoder complexity increasing and memory usage in this paper.

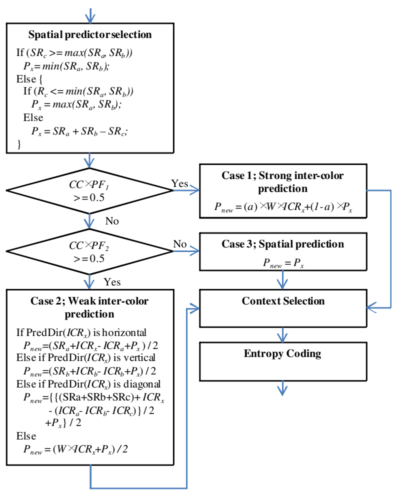

Fig. 3. Flowchart of spatio-color prediction

Finally, in this paper, the same prediction scheme like JPEG-LS is exploited on case 3 which is considered no correlation among the other color planes.

The proposed spatio-color algorithm can be summarized as the flowchart shown in Fig.3

3 Context Modeling

The context model exploited on JPEG-LS can be described by two following steps, calculating the difference of neighboring pixels at first and then quantizing into 9 regions through the thresholds of [*-T3, -T2, -T1, 0, T1, T2, T3*].

The obtainable differences from 4 neighboring pixels are three kinds as $D_1=SR_d-SR_b$, $D_2=SR_b-SR_c$, $D_3=SR_c-SR_a$ and total 729 contexts (9×9×9=729) are exploited. However, practically, 364 contexts are employed in regular mode because the same context is used for $[D_1, D_2, D_3]$ and $[-D_1, -D_2, -D_3]$ simultaneously and 2 more contexts are used for run mode additionally.

A new context model is needed for 4K image because the conventional context model does not fully reflect the considerably high spatial correlation. But the prediction scheme of G color plane is the same as conventional method of JPEG-LS so that the context model exploited in G color plane would face a corruption when the conventional model is injured. A new method for maximization of spatial correlation keeping conventional method free from corruption should be invented. So the introduced method in this paper is to classify the average residual of neighboring pixels in to 3 categories at first and then to decide an actual context number of the current pixel after calculating the difference of neighboring pixels as in JPEG-LS as per pseudo code shown in fig. 4.

The proposed method in this paper has been employed under the assumption that spatial correlation has not been perfectly removed, which means the current pixel would have similarity though the direction of spatial prediction is different from that of neighboring pixels. Accordingly, the remained residual might have similar level. For that reason, classifying a similar level residual into the same context would bring improvement of coding efficiency in entropy coding. And the actual context number can be selected within the category as the same method in JPEG-LS because each category has 364 contexts respectively.

1. Calculate the average of ($SR_a \sim SR_d$) residual
   (AVG(SRa ~ SRd))
2. If $AVG(SR_a \sim SR_d) > 0x08$
   Context Number = High context (732 <= ContextNum)
3. Else if $AVG(SR_a \sim SR_d) > 0x03$
   Context Number = Middle context (367 <= ContextNum < 732)
4. Else
   Context Number = Low context (ContextNum < 365)

Fig. 4. Flowchart of spatio-color prediction

The context model using the level of residual can be employed in inter-color prediction. As residual is classified by the level of residual, the improved coding performance is expected if some of similar residual can be grouped together even tough difference prediction is applied.

4 Test Sequences

The test images used in this paper is from the courtesy of Korea Film Council (KEFIC) [8] taken by RED ONE™ camera of RED Digital Cinema Camera Company [9].

Each image consists of 4096×1716 resolution with RGB 4:4:4 12bit depth and various scenes are taken. The stored file format is DPX and images have been manipulated linear tone mapping to change 12bit depth into 10bit depth for this paper.

The test images which were used in simulation were made to compose similar continuous scenes up to 60 frames and total output bitstream size of whole frames was compared. The comparison of total bitstream size was applied mainly because of the expectation that the developed codec works like high speed real-time video codec in UHDTV system. Thus this research is purposed to measure the coding gain of whole sequence not to measure that of one frame, however during the simulation it is observed that coding gain of similar scene images is not so much fluctuated.

We would like to emphasize that no temporal correlation algorithm has been exploited and more emphasize that each image can be divided separately and can be edited freely as per frame by frame even tough encoded at once.

To give better understanding for test images, short description is given as following. Each image is consisted as a set of scenes and each set shows both the first image and the last image respectively.

Fig. 5. The first image of test image sequence 1 (upper) and the last image of test image sequence 1 (lower)

The first sequence as shown in fig.5 shows not much difference between the first image and the last image, but some motions can be found at the movement of big wheel on left side and the movement of windmill on right side.

Fig. 6. The first image of test image sequence 2 (upper) and the last image of test image sequence 2 (lower)

In the second sequence as shown in fig.6, the first image starts with flying helicopter on the left top and ends with scene changes as shown in fig.6 upper image. The fast motion of rollercoaster needs to draw attention on the second sequence. The complexity of image is quite high and motion is also fast.

Fig. 7. The first image of test image sequence 3 (upper) and the last image of test image sequence 3 (lower)

The third sequence as shown in fig.7 shows a parade at an amusement complex and an actress in the center moves very rapidly her arms and the mask on top of stick moves top to down continuously and scene change with zoom-in is occurred as the last image as shown fig.7 lower image. This sequence contains various colors and fast motion of foreground objects.

Fig. 8. The first image of test image sequence 4 (upper) and the last image of test image sequence 4 (lower)

The fourth sequence as shown Fig.8 shows a pair of lovers who sits on the back of a camel and moves forward slowly. The large object moves continuously and slowly and the object moves gradually from bright region to dark region.

Fig. 9. The first image of test image sequence 5 (upper) and the last image of test image sequence 5 (lower)

In the last sequence, the fifth one, there is continuous camera panning on left top direction with showing continuous water streams and water fall.

The tested images used in this paper cover various kinds of scene including almost no movement, scene change, camera panning, small fast object, and large slow object, in which they contains all possible scenes that UHDTV system may encounter.

All images employed are tested under developed UHDTV system, which requires 60 frames per second but only 55 frames were used on the fourth set in order to eliminate scene change.

5 Simulation Results

All test images used for the simulation have 4096×1716 resolution and RGB 4:4:4 10bit depth as described section 4 and consecutive 60 frames are coded excluding fourth test images. In the developed UHDTV system, the maximum supported frame rate will be 60 frames per second and the proposed algorithms shall be verified during 1 second at least. The total size of output bitstream including header syntax was compared to that of JPEG-LS as shown Table. 1. The proposed algorithms are based on lossless image coding, and the reconstruction image is perfectly identical to input original image so that no PSNR comparison is prepared.

Table 1. Font sizes of headings. Table captions should always be positioned *above* the tables

Image	Sequence Number	Total Frame Number	JPEG-LS (Byte)	The Proposed (Byte)	Gain (%)
	Test 1	60 Frames	621,097,483	587,798,614	5.36
	Test 2	60 Frames	579,323,321	546,083,094	5.74
	Test 3	60 Frames	581,731,982	553,109,207	4.92
	Test 4	55 Frames	662,060,078	629,683,219	4.89
	Test 5	60 Frames	784,091,431	732,239,196	6.61
	Average				5.50

As shown in Table.1, the proposed algorithms shows up to 6.6%~4.9% coding gain in terms of output bitstream compared to JPEG-LS and 5.5% coding gain can be achieved in average.

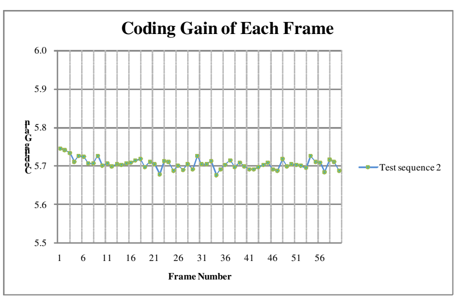

Fig. 10. Coding gain of each frame is compared to JPEG-LS on test sequence 2

In case of the second sequence and the third sequence which contain scene change in the middle of them however, during the simulation, the scene change shows no performance degradation because each frame is coded independently. Fig.10 shows coding gain of each frame compared to JPEG-LS.

6 Conclusion

In this paper, the new lossless algorithm called spatio-color prediction which is very suitable for large image exceeding 2K is proposed and the new context modeling based on residual level is also proposed. The proposed algorithms show improved coding gain up to 6.6% ~ 4.9% in terms of output bitstream size and 5.5% coding gain in average in comparison to JPEG-LS.

The proposed algorithm is verified by the simulation according to various images exploited, which contains various camera actions, foreground movement, and color. Also, the proposed algorithms are adapted in the developed UHDTV system and utilized real-time codec practically.

References

1. ISO/IEC JTCl/SCPS WGl (JPEG/JBIG): Information Technology - Lossless and Near-Lossless Compression of Continuous-Tone Still Images, Draft International Standard DIS14495-1 (JPEG-LS) (1998)
2. Weinberger, M.J., Seroussi, G., Sapiro, G.: LOCO-I: A Low Complexity, Context-Based, Lossless Image Compression Algorithm. In: Proc. of the 1996 Data Compression Conference, pp. 140–149 (1996)
3. Wu, X.: An Algorithmic Study on Lossless Image Compression. In: Proc. of the 1996 Data Compression Conference, pp. 150–159 (1996)
4. Wu, X.: Lossless Compression of Continuous-Tone Images via Context Selection, Quantization and Modeling. IEEE Trans. on Image Processing 6(5), 656–664 (1997)
5. Wu, X., Memon, N., Sayood, K.: A Context-Based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for continuous-Tone Images. ISO/IEC JTC l/SC 29/WC 1 document No. 202 (1995)
6. Wu, X., Choi, W.K., Memon, N.: Lossless Interframe Image Compression via Context Modeling. In: Proc. of the 1998 Data Compression Conference, pp. 378–387 (1998)
7. Barequet, R., Feder, M.: SICLIC: a simple inter-color lossless image coder. In: Proc. of the 1999 Data Compression Conference, pp. 501–510 (1999)
8. Korea Film Council, `http://koreanfilm.or.kr/`
9. RED Digital Cinema Camera Company, `http://www.red.com/cameras`
10. Matsuda, I., Kaneko, T., Minezawa, A., Itoh, S.: Lossless Coding of Color Images using Block-Adaptive Inter-Color Prediction. In: IEEE International Conference on Image Processing 2007, vol. 2, pp. II-329–II-332 (2007)
11. Fukuma, S., Iwahashi, M., Kambayashi, N.: Lossless Color Coordinate Transformer for Lossless Color Image Coding. In: Proc. of IEEE Asia-Pacific Conf. on Circuits and Systems, pp. 595–598 (1998)
12. Golchin, F., Paliwal, K.K.: Minimun-entropy clustering and its applcation to lossless image coding. In: Proc. of the IEEE International Conference on Image Processing (1997)
13. Memon, N.D., Sayood, K.: Lossless compression of video sequences. IEEE Trans. Commun. 44(10), 1340–1345 (1996)