Region-based Spectral Correlation Estimator for Color Image Coding

Noyoon Kwak*

Abstract

This paper is related to the Region-based Spectral Correlation Estimation(RSCE) coding method that makes it possible to achieve the high-compression ratio by estimating color component images from luminance image. The proposed method is composed of three steps. First, Y/C bit-plane summation image is defined using normalized chrominance summation image and luminance image, and then the Y/C bit-plane summation image is segmented for extracting the shape information of the regions. Secondly, the scale factor and the offset factor minimizing the approximation square errors between luminance image and R, B images by the each region are calculated. Finally, the scale factor and the offset factor for the each region are encoded into bit stream. Referring to the results of computer simulation, the proposed method provides more than two or three times higher compression ratio than JPEG/Baseline or JPEG2000/EBCOT algorithm in terms of bpp needed for encoding two color component images with the same PSNR.

Keywords : Color Image Coding, Spectral Correlation Estimator, Image Segmentation, Image Compression

컬러 영상 부호화를 위한 영역 기반 스펙트럴 상관 추정기

곽노윤*

요약

본 논문은 휘도 성분 영상으로부터 컬러 성분 영상들을 추정 부호화함으로써 높은 압축률을 달성할 수 있는 영역 기반 스펙트럴 상관 추정 부호화 방법에 관한 것이다. 제안된 방법은 3단계로 구성되어 있다. 우선, 정규 색차합 영상과 휘도 영상을 이용해 산출한 Y/C 비트 평면합 영상을 대상으로 영상 분 할을 수행하여 영역들의 형상 정보를 추출한다. 이후, 각 영역 단위로 휘도 영상과 R, B 영상 간의 근사 화 자승 오차가 최소가 되도록 하는 비례 인자와 가감 인자를 산출한다. 최종적으로, 각 영역의 비례 인 자와 가감 인자를 비트스트림으로 부호화한다. 컴퓨터 시뮬레이션 결과에 따르면, 제안된 방법은 동일한 PSNR에서 두 칼라 성분 영상을 부호화하기 위해 소요되는 bpp를 비교할 때, JPEG/베이스라인 혹은 JEPG2000/EBCOT 알고리즘에 비해 2배 혹은 3배 이상의 압축률을 제공한다.

키워드 : 컬러 영상 부호화, 스펙트럴 상관 추정기, 영상 분할, 영상 압축

1. Introduction

- Corresponding Author : Noyoon Kwak
 Received: December 15, 2016
 Revised: December 26, 2016
 Accepted: December 31, 2016
 * Division of Information and Communication
 Engineering, Baekseok University
- Tel: +82-41-550-0505, Fax: +82-41-550-9083 email: nykwak@bu.ac.kr
- This Work was financially supported by 2016 research grant from Baekseok University

It is a matter of common knowledge that chrominance signal is less sensitive to the eyes than luminance signal and, that there exists high spectral correlation between color component images. Especially, the spectral correlation can easily observed in the spatial domain through the human eye, without the mathematical analysis through the color coordinate transform. And many methods are widely studied to effectively remove the spectral redundancy [1][2].

One key problem in image compression is to analyze the statistical characteristics of source data for better entropy coding [1]. After devoting many efforts in this area in the last decade researchers have come up with several elegant ways for modeling the statistical characteristics. For example, in the embedded zerotree wavelet (EZW) coder [3], coefficients after discrete wavelet transform (DWT) can be represented by a zerotree structure. In the set partitioning in hierarchical trees (SPIHIT) coder [4] a so-called spatial orientation tree (SOT) was adopted. And in JPEG2000 [5][6], the values of neighboring pixels are used as context when modeling the statistics of a current pixel.

The above methods were designed and optimized for monochrome images. However, most digital images are colorized. When coding color images each component is usually coded independently without any inter-color prediction [6][7]. Although being extensions of algorithms for coding monochrome images can simplify coding system design, coding sacrificed efficiency will be since the correlation among different color components is not exploited. In fact, one can often easily recognize the shapes of the chrominance components from that of the luminance component when they are separately displayed. The phenomenon indicates that most chrominance changes in a real scene are accompanied by a luminance change [8]. Having noticed the inter-color correlation different color components, among some researchers began placing their attention this field For instance color-EZW(CEZW) [9] and color-SPIHT(CSPIHT) [10] were proposed to exploit underlying inter-color correlation by expanding the existing zerotree or SOT structure across the spectral planes. By using the inter-color correlation, these methods obtain а better performance than the traditional coding schemes. Though there exists the inter-color correlation, it does not always mean that a significant chrominance change will be accompanied by a significant luminance change at the same pixel position. Moreover, the correspondence becomes even more random after DWT. Generally speaking, the significant wavelet coefficients generated by the chrominance changes may be found within a neighboring region of the luminance coefficients. In such a situation, the simple one-one mapping adopted in CEZW and CSPIHT is not always effective.

In the international standards like MPEG-1, MPEG-2, JPEG, JPEG2000, FCC/HDTV, H.261, H.263, H.264, etc., R, G, B are converted to Y, Cb, Cr which have spectrally low correlation and then Chrominance signals Cb, Cr are down-sampled with 4:1:1, 4:2:2 etc. for compressing these signals in the low resolution[11][12]. Through the above kinds of processes, the spectral correlation between the color components is indirectly removed. In this case, each color component is individually compressed and encoded with some loss through the frequency transform like Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT). In spite of these process, there is a problem that much of the spectral correlation is still remained because, based on the characteristic of the human visual system in which high frequency color component is less sensitive, these processes are to reduce each color signal individually through the analysis of the low frequency. And now the more advanced image coding method should be needed that can remove the correlation more efficiently and consequently offers higher compression ratio to match the demand of extremely high compression ratio.

We have proposed BSCE(Block-based Spectral Correlation Estimation) coding method [13] that can estimate another component image from one component image by the block, which is basically exploiting the spectral correlation existed between the color component images. Because the BSCE method is estimating the spectral correlation and directly remove it, the method can offer higher compression ratio comparing to the color image compression method which remove the spectral correlation indirectly through the low frequency analysis. But this method results in the serious blocking effect in the local region which has the low color correlation between the reference image and color component image. And dividing the block for estimation into the smaller to reduce the effect can cause problem that the the transmission load increases at the higher rate than the improvement performance does. In the proposed paper, we take the segmentation step to extract the shape information for each region, and determine the scale factor and offset factor for each region that make the approximation square error between the luminance image and R, B images minimum. Finally, we propose the new approach of region-based color image coding method from encoding the resulting factors.

2. Proposed RSCE Coding

The proposed RSCE(Region-based Spectral Correlation Estimation) coding method is composed of the two stages. (Figure 1) is shown an example of the proposed RSCE coder and decoder. The first stage is for the extraction of shape information for each region which is basic unit to estimate the spectral correlation, and the second stage is for referencing the shape information for extracted region and finding the optimum scale factor and offset factor for each region.



(Figure 1) An Example of RSCE Coder and Decoder

2.1 Extracting Shape Information

In the proposed RSCE method, we use block-based image segmentation algorithm [14][15] which is known for effectively restricting the over-segmentation as well as reducing the amount of calculation to design new block-based image segmentation method that can effectively extract the shape information for the region.

First, normalized chrominance summation image is obtained by normalizing the image which is summed up the absolutes of color-differential values between R, G, Bimages. Secondly, upper 2 bits of the luminance image and upper 6bits of the normalized chrominance summation image are bitwise operated by the pixel to generate the Y/C bit-plane summation image. Next, the Y/C bit-plane summation image divided into predetermined block size, is classified into monotone blocks, texture blocks and edge blocks, and then each classified block is merged to the regions including one more blocks in the individual block type, and each region is selectively allocated to unique marker according to predetermined marker allocation rules. Finally, fine segmented results are

obtained by applying the watershed algorithm to each pixel in the unmarked blocks.

(Figure 2) is shown the final results of the segmentation overlapped onto the luminance image for the LENA. As shown in computer simulation results, the proposed method offers reasonable segmentation results in edge regions with lower contrast owing to the regional characteristics of the color components reflected in the Y/C bit-plane summation image.



(a) LENA (b) SAILBOAT (Figure 2) Segmented Image

2.2. RSCE Coding

2.2.1 The Estimation of Region-based Spectral Correlation

To implement the region-based spectral correlation estimator, first, we extract the luminance image Y from the input R, G, B images to use it for the reference image, and the scale factor and offset factor used to replace the R image and B image. The G image is set for the reference image, but since it is more sensitive to the characteristic of the human visual system as well as it does not represent the spectral correlation between the color component images better than Y image, Y image is selected for the reference image.

Equation (1) expresses the set of pixels of the Y, B, and R images, where M is the number of the vertical pixels and N is the number of the horizontal pixels of the each image.



(Figure 3) Region-based Spectral Correlation Estimation

$$Y = \bigcup_{i=1}^{N} \bigcup_{j=1}^{M} y(i \times M + j)$$

$$R = \bigcup_{i=1}^{N} \bigcup_{j=1}^{M} r(i \times M + j)$$

$$B = \bigcup_{i=1}^{N} \bigcup_{j=1}^{M} b(i \times M + j)$$
(1)

$$\begin{split} \boldsymbol{R}_{i\pi}(k) &= \sum_{i=1}^{con[k]} (S_{\pi}(k) y(v[i] \times N + h[i])) \\ &+ O_{\pi}(k) - r(v[i] \times N + h[i]))^{2} \\ \boldsymbol{R}_{i\pi}(k) &= \sum_{i=1}^{con[k]} (S_{\pi}(k) y(v[i] \times N + h[i])) \\ &+ O_{\pi}(k) - b(v[i] \times N + h[i]))^{2} \end{split}$$
(2)

The approximation square error of R the and B images for the Y image is defined by Equation (2) to determine the scale factor and offset factor which approximates the R and Bimages so that they have least squared distance from the pixel set of the Y image. h[i] and v[i] are the vertical and horizontal coordinates of the pixel contained in an object region with marker k, and cnt[k] is the number of the pixel in that object region. $S(k)_{B,O}(k)_{B}$ is the scale factor and offset factor of the *B* image, and $S(k)_R, O(k)_R$ is the scale factor and offset factor of the *R* image. And then, after applying the partial derivative with the scale factor and offset factor as variables to approximation square error, we find the scale factor and offset factor that make the approximation square error minimum from the partial derivative equation assigned to be zero. Equation (3) represents the process.

$$\frac{\partial R_{YR}(m,n)}{\partial S_R(m,n)} = 0, \qquad \frac{\partial R_{YR}(m,n)}{\partial O_R(m,n)} = 0$$
$$\frac{\partial R_{YB}(m,n)}{\partial S_B(m,n)} = 0, \qquad \frac{\partial R_{YB}(m,n)}{\partial O_B(m,n)} = 0$$
(3)

Equation (4) is the result for the scale factor which meets the condition of Equation (3).

$$\begin{split} S_{R}(k) &= (cnt [k] \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i])r(v[i] \times M + h[i]) \\ &- \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i]) \sum_{i=1}^{cot[k]} r(v[i] \times M + h[i])) \\ / (cnt [k] \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i])^{2} - (\sum_{i=0}^{cot[k]} y(v[i] \times M + h[i]))^{2}) \\ S_{B}(k) &= (cnt [k] \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i])b(v[i] \times M + h[i]) \\ &- \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i]) \sum_{i=1}^{cot[k]} b(v[i] \times M + h[i])) \\ / (cnt [k] \sum_{i=1}^{cot[k]} y(v[i] \times M + h[i])^{2} - (\sum_{i=0}^{cot[k]} y(v[i] \times M + h[i]))^{2}) \end{split}$$
(4)

The resulting equation for the scale factor which meets the condition of Equation (3) is the same as Equation (5).

$$O_{R}(k) = \frac{1}{cnt[k]} \left(\sum_{i=1}^{cnt[k]} r(v[i] \times M + h[i]) - S_{R}(k) \sum_{i=1}^{cnt[k]} v(v[i] \times M + h[i]) \right)$$

$$O_{B}(k) = \frac{1}{cnt[k]} \left(\sum_{i=1}^{cnt[k]} b(v[i] \times M + h[i]) - S_{B}(k) \sum_{i=1}^{cnt[k]} v(v[i] \times M + h[i]) \right)$$
(5)

If the denominator of the $S_R(k)$ and $S_B(k)$

is the same as Equation (6), that is, the deviation(σ_k^2) of the luminance for the object is the same as the zero, the calculation error can possibly occur.

$$cnt[k] \sum_{i=1}^{cnt[k]} y(v[i] \times M + h[i])^{2} - \left(\sum_{i=1}^{cnt[k]} y(v[i] \times M + h[i])\right)^{2} = cnt[k]^{2} \sigma_{k}^{2} = 0$$
(6)

So in that cases, we solve the problem by appling Equations (7) to R and B images.

$$S_{R}(k) = 0$$

$$O_{R}(k) = \frac{1}{cnt[k]} \sum_{i=1}^{cnt[k]} r(v[i] \times M + h[i])$$

$$S_{B}(k) = 0$$

$$O_{B}(k) = \frac{1}{cnt[k]} \sum_{i=1}^{cnt[k]} b(v[i] \times M + h[i]) + O_{R}(k)$$
(7)

Equations (8) and (9) represent the Mean Square Error(MSE) of the estimated R and B images.

$$R_{MSE} = \frac{1}{cnt[k]} \left[\sum_{i=1}^{cnt[k]} b(v[i] \times N + h[i])^2 + S_R \left(S_R \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i])^2 - \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i]) \cdot b(v[i] \times N + h[i]) + 2O_R \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i]) \right) + O_R \left(cnt[k] \cdot O_R - 2 \sum_{i=1}^{cnt[k]} b(v[i] \times N + h[i]) \right) \right]$$
(8)

$$B_{MSE} = \frac{1}{cnt[k]} \left[\sum_{i=1}^{cnt[k]} r(v[i] \times N + h[i])^2 + S_B \left(S_B \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i])^2 - \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i]) \cdot r(v[i] \times N + h[i]) + 2O_B \sum_{i=1}^{cnt[k]} y(v[i] \times N + h[i]) \right) + O_B \left(cnt[k] \cdot O_B - 2 \sum_{i=1}^{cnt[k]} r(v[i] \times N + h[i]) \right) \right]$$
(9)

2.2.2 The Compensation of Region-based Spectral Correlation

After the variable length decoding for the received bit-stream, luminance image is decoded through the texture decoding. The decoded R, B images are obtained by performing the compensation decoding for each object repeatedly with Equations (10) and (11). K is the number of the marker in the given image, the number of the object.

$$R = \bigcup_{k=1}^{K} r(k)$$

$$r(k) = \bigcup_{k=1}^{ort[k]} (S_R(k) \times y(v[i] \times N + h[i]) + O_B(k)$$
(10)

$$B = \bigcup_{k=1}^{K} b(k)$$

$$b(k) = \bigcup_{k=1}^{cn[k]} (S_B(k) \times y(v[i] \times N + h[i]) + O_R(k)$$
(11)

When completing the above process, the other color component image, G image, can be easily decoded with Equation (12).

$$G = \bigcup_{k=1}^{K} g(k)$$
$$g(k) = \frac{1}{0.587} (y(k) - 0.299r(k) - 0.144b(k))$$
(12)

3. Simulation and Results

To verify the universal validity of the proposed RSCE coding method, we used JPEG/Baseline algorithm and JPEG2000/ comparison EBCOT algorithm for and performed computer simulation with Microsoft Visual C++. From the simulation result for the the PSNR hit rate at same with TABLETENNIS(#50-#64) sequence, we recognize that bit per pixel to maintain 26dB of the average PSNR for R image and B image and 29dB of the decoding performance for Y image are 0.0487, 0.0322, and 0.0122 bpp of the average bit rate for JPEG/Baseline algorithm and JPEG2000/EBCOT algorithm and RSCE coding method, respectively. So this

RSCE method has about 4.0 times and 2.64 times higher compression ratio than JPEG and EBCOT, respectively.

So this RSCE method has about 4.0 times and 2.64 times higher compression ratio than JPEG and EBCOT, respectively. (Figure 4) shows the decoding results with these different methods for the 62nd frame of TABLETENNIS sequence under the same condition. As shown in (Figure 4), the proposed RSCE coding method offers better subjective performance comparing to the JPEG and EBCOT.



(a) JPEG/baseline Method



(b) JPEG2000/EBCOT Method



(c) Proposed RSCE coding method (Figure 4) Decoding Results with Different Methods for TABLETENNIS #62(Y=29dB, R+B/2=26dB)



Because basically, this method reproduces the R and B components with the basis of the shape information of the region, the contour of the region can be well preserved even if the quality of the Y image is seriously damaged. When we evaluate the compression ratio quantitatively of the R and B image, proposed method offers at least 2 times higher ratio additionally than that of the other image standards.

(Figure 6) shows average PSNR of the R and B images with changing the PSNR of Y image while the bit per pixel of the Cb, Cr or R, B images are fixed to the certain value constantly. The proposed RSCE coding method provides more than 3 times higher compression ratio as well as has the property that it is less sensitive to the quality of the Y image. Especially, even in the position that the quality of the Y image is seriously damaged, it serves to decrease the quality smoothly comparing to the EBCOT. SNR comparing to the EBCOT when the PSNR of the Y image increases at a certain cross-point, but this is almost nothing comparing to the improvement of the performanceoccurred when the PSNR of Y image decreases. Especially, the this property is appeared more clearly and more positively in rather complex images, like LENA. SAILBOAT. MILKCROWN, and

AIRPLANE. Also, it offers more than about 5 times higher compression ratio than EBCOT as well as maintains the increasing and decreasing characteristics at the certain vertex in the graph. From these properties, it is expected that the proposed RSCE coding method hasadvantages for the data transmission in the very low bit rate channel.





(Figure 6) Performance comparison of the results in PSNR and bpp

(Figure 7) shows compression ratio of the RSCE coding method for the total of R, G, B data when the visual qualities of the Rand B images are set to be equal to that of EBCOT.



(Figure 7) Comparison of the compression ratio for LENA image

The contribution for the whole compression ratio of the color as compression ratio additionally, is image increases component little that the compression very in the case comparatively, ratio is low but the contribution increases at the high rate as it goes to the high compression ratio. So as mentioned above, it is apparent that the proposed RSCE coding method provides the prominent performance in the applications of the image CODEC for the very low bit rate transmission.

4. Conclusions

As we can see in the computer simulation, the proposed method offers excellent performance comparing to the EBCOT in JPEG2000/EBCOT and JPEG/Baseline algorithm. Also, while these early methods provide degraded results in the contour and high frequency region, proposed method is evaluated to have better performance in the subjective quality because it reproduces the contour of region almost perfectly which isvery sensitive to the human visual system. Especially, the proposed method either uses the shape information directly offered to encode the Y image or extracts the shape information from decoded Y image because the method has the segmentation unit with the same algorithm both in the encoder and the decoder.

So there is no need to transmit the shape additionally information data to encode estimated color component. This makes it possible for the proposed method to be used with both region based image CODEC and the spectral CODEC of the block or frame based image CODEC comfortably. And the spectral correlation mainly handled in the proposed RSCE codingmethod is expected to have various applications in the MPEG-7 standardization process, for example, for the descriptor for image retrieval being studied as well as in the image compression.

References

- Xing San, Hua Cai, and Jiang Li, "Color Image Codin g by using Inter-Color Correlation," in Proc. of IEE E International Conference on Image Processing, pp. 3117–3120, Oct. 2006.
- [2] Y. Roterman, M. Porat, "Progressive Image Coding Using Regional Color Correlation," in Proc. of 4th

EURASIP Conference focused on Video/Image Proc essing and Multimedia Communications, vol. 1, pp. 65–70, Jul. 2003.

- [3] J. M. Shapiro, "Embedded Image Coding Using Zerot rees of Wavelets Coefficients," IEEE Transactions on Signal Processing, vol. 41, pp. 3445–3462, Dec. 1993.
- [4] A. Said and W. A. Pearlman, "A New Fast and Effici ent Image Codec Based on Set Partitioning in Hierar chical Trees," IEEE Transactions on Circuits and System for Video Technology, vol. 6, pp. 243–250, Jun. 1996.
- [5] Majid Rabbani and Rajan Joshi, "An overview of the JPEG 2000 still image compression standard," Signa l Processing: Image Communication, vol. 17, pp. 3–4 8, 2002.
- [6] ISO/IEC 15444-l, "Information Technology-JPEG20 00 -Image Coding System-Part 1: Core Coding Syst em," 2000.
- [7] G. K. Wallace, "The JPEG Still Picture Compression Standard," Comm. of the ACM, vol. 34, pp. 30–44, Apr. 1991.
- [8] J. O. Limb and C. B. Rubinstein, "Plateau Coding of the Chrominance Component of Color Picture Sig nals," IEEE Transactions on Communications, vol. 22, pp. 812–820, Jun. 1974.
- [9] K. Shen and E. J. Delp, "Color Image Compression Using an Embedded Rate Scalable Approach," in Pr oc. of IEEE International Conference on Image Proc essing, Santa Barbara, 1997.
- [10] A. A. Kassim and W. S. Lee, "Embedded Color Imag e Coding Using SPIHT with Partially Linked Spatial Orientation Trees," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 203– 206, Feb. 2003.
- [11] David Taubman, "High Performance Scalable Imag e Compression with EBCOT," IEEE Transaction on

Image Processing, Mar. 1999.

- [12] T. L. Lin, P. S. Liaw, S. L. Chen, C. H. Hsia, and S. Y. Lin, "Chroma 4:2:2 Subsampling for Bayer Pat tern in H.264 Video Coding," IET Electronics Letter s, pp. 608–609, vol. 52, no. 8, Apr. 2016.
- [13] Noyoon Kwak, Min-Su Shin, Dae-Gwon Jeong, and Byong-Won Hwang, "A Color Image Coding by Esi mating Spectral Correlation Based on Wavelet Tran sform," Journal of The Institute of Electronics Engin eerers of Korea, vol. 37– SP, no. 1, pp.49–58, Jan. 2000.
- [14] C. S. Won, "Extracting and Homogeneous Regions Using Block-based MAP Segmentation," in Proc. of WIAMIS 97, pp. 162–172, 1997.
- [15] Jae-Hyuk Lim, Dong-Kwon Park, and Chee-Sun Won, "The Image Segmentation Algorithm with Blo ck Classification and Watershed," Journal of the Inst itute of Electronics Engineers of Korea, vol. 36–S, no. 1, pp.81–89, Jan. 1999.



Noyoon Kwak

1994 : Korea Aerospace University, Avionics Engineering (B.S.)
1996 : Korea Aerospace University, Avionics Engineering (M.S.)
2000 : Korea Aerospace University, Avionics Engineering (Ph.D.)

2000~now : Professor in Backseok University, Korea Research Interests : Video Coding, VR/AR, Object Tracking, Image Processing and Computer Vision, etc.