PAPER

Colorization Based Image Coding by Using Local Correlation between Luminance and Chrominance

Yoshitaka INOUE[†a)], ***Student Member***, **Takamichi MIYATA**[†], ***and*** **Yoshinori SAKAI**[†], ***Members***

SUMMARY Recently, a novel approach to color image compression based on colorization has been presented. The conventional method for colorization-based image coding tends to lose the local oscillation of chrominance components that the original images had. A large number of color assignments is required to restore these oscillations. On the other hand, previous studies suggest that an oscillation of a chrominance component correlates with the oscillation of a corresponding luminance component. In this paper, we propose a new colorization-based image coding method that utilizes the local correlation between texture components of luminance and chrominance. These texture components are obtained by a total variation regularized energy minimization method. The local correlation relationships are approximated by linear functions, and their coefficients are extracted by an optimization method. This key idea enables us to represent the oscillations of chrominance components by using only a few pieces of information. Experimental results showed that our method can restore the local oscillation and code images more efficiently than the conventional method, JPEG, or JPEG2000 at a high compression rate.

key words: *image coding, colorization, total variation, correlation between luminance and chrominance*

1. Introduction

Colorization [1], [2] is a process that restores a complete color from a given complete gray-scale image and incomplete color information. This method is based on the assumption that neighboring pixels that have similar luminance also have similar chrominance. Recently, a novel approach to color image compression based on colorization has been presented [3]–[6]. The conventional method [3] reduces the coding rate by transforming whole chrominance components into small color assignments at the encoder. The chrominance components are restored by propagating the color assignment by a colorization [1] at the decoder. Although the conventional method outperforms JPEG from the viewpoint of its visual quality, the decoded chrominance components tend to lose the local oscillation that the original images had. Additional color assignments are required to restore these oscillations precisely, but these decrease the coding efficiency.

On the other hand, previous studies [7]–[12] suggest that an oscillation of chrominance components correlates with the oscillation of the corresponding luminance component. In this paper, we focus on the correlation that exists between luminance and chrominance in a local area of images and present a new colorization-based coding method that can restore oscillations of chrominance. First, the input image is separated into the texture (oscillations) and the geometry (smooth areas and edges) components by using a total variation regularization method [13], [14]. Second, the geometry component is divided into a few clusters to group the grid points (vertices). Note that the location and grouping of these vertices can be shared at both the encoder and decoder side without any overhead. For texture components, we assume that the local correlation of oscillation is approximated by a linear function at each cluster, and we obtained the chrominance (for geometry) and the coefficient of the linear function (for texture) that minimize the decoded error at each cluster. The luminance component is compressed by JPEG2000. The cluster reference information of the vertices and the chrominance and the coefficient of each cluster are sent to the decoder. Our method enables significant rate reduction by sharing the same vertices at both the encoder and decoder sides. At the decoder, the chrominance geometry components are decoded from color vertices on the luminance image, and the chrominance texture components are restored by the coefficients.

The paper is organized as follows. First, we briefly introduce the existing colorization method [1] in Sect. 2. In Sect. 3, we describe some colorization-based image coding methods from our previous work. In Sects. 4 and 5, we introduce related works on geometry-texture separation and inter-color correlation. We present a new colorization-based image coding method based on correlation between luminance and chrominance in texture component in Sect. 6. Experimental results of our method are provided in Sect. 7. Finally, Sect. 8 concludes this paper.

2. Colorization

The colorization method developed by Levin et al. [1] is used to restore natural chrominance components by estimation from small number of pixels that have chrominance value. The method is based on the assumption that neighboring pixels that have similar luminance also have similar chrominance.

Luminance similarity of two pixels $\boldsymbol{r}$ and $\boldsymbol{s}$ is defined as

$$w_{\boldsymbol{rs}} = \begin{cases} e^{-\{Y(\boldsymbol{r})-Y(\boldsymbol{s})\}^2/2\sigma_{\boldsymbol{r}}^2} & (\boldsymbol{s} \in N(\boldsymbol{r})) \\ 0 & (\boldsymbol{s} \notin N(\boldsymbol{r})). \end{cases} \quad (1)$$

Manuscript received March 23, 2011.
Manuscript revised August 20, 2011.

[†]The authors are with Communications & Integrated Systems, Graduate School of Science and Engineering, Tokyo Institute of Technology, Tokyo, 152–8552 Japan.
a) E-mail: inoue@net.ss.titech.ac.jp

DOI: 10.1587/transinf.E95.D.247

Where, $Y(\boldsymbol{r})$ represents the luminance at $\boldsymbol{r}$, and $\sigma_{\boldsymbol{r}}^2$ is variance around $\boldsymbol{r}$, $N(\boldsymbol{r})$. $N(\boldsymbol{r})$ is defined as eight neighboring pixels of $\boldsymbol{r}$ in [1]. Describing the number of pixels of an input image as n, a_{ij} that represents i-th row and j-th column factor of matrix $A \in \mathbb{R}^{n \times n}$ is defined as follows.

$$a_{ij} = \begin{cases} 1 & (\boldsymbol{i} = \boldsymbol{j}) \\ 0 & ((\boldsymbol{i} \neq \boldsymbol{j}) \wedge (\boldsymbol{i} \in V)) \\ -w_{\boldsymbol{ij}} / \Sigma_{s \in N(\boldsymbol{i})} w_{\boldsymbol{is}} & (otherwise). \end{cases} \quad . \tag{2}$$

Here, $\boldsymbol{i}$ and $\boldsymbol{j}$ are pixels corresponding to i and j respectively, and V is a set of given chrominance pixels. Describing a column vector $Q \in \mathbb{R}^n$ whose factor has a chrominance value if V has a corresponding pixel, or 0, and restored chrominance vector $X \in \mathbb{R}^n$, Colorization [1] is formulated as

$$X = A^{-1}Q. \tag{3}$$

X is restored by propagating Q along with segments of luminance component.

3. Colorization-Based Coding

Recently, color image compression by using colorization at decoder has been presented [3]–[6].

Conventional colorization-based coding [3] represents chrominance components of all pixels as a few color lines at the encoder. At the decoder, chrominance components are recovered by using a colorization method [1]. However, the decoded chrominance components lose the local oscillation that the original images had.

Another method [4] extracts some color vertices at the encoder. Color vertices are propagated along with areas segmented by luminance edges [2]. However, if the number of vertices is small in the areas that have dense oscillations of luminance, chrominances cannot be propagated correctly. Moreover, chrominance components are radically smoothed, similar to those in the work of Levin et al. [1].

Since the local oscillations represent the textures and shadows of materials, the loss of such oscillations leads subjective quality to considerably degrade. A large number of color assignments such as the vertices or the lines are required to restore these oscillations precisely. However, the increase in the number of color assignments makes the compressed data size large.

An alternative colorization-based image coding method to restore oscillations has been proposed [5]. Coefficients that describe the correlation between luminance and chrominance in the texture component are used to restore oscillations [5]. However, achieving drastic bit-rate reduction is difficult, since the coordinates information of lines must be transmitted for color assignment.

4. Geometry-Texture Separation

Image denoising methods that use total variation (TV) regularization have been proposed by Rudin et al. [13]. Let $\boldsymbol{x} \in \mathbb{R}^2$ be a coordinate in the image, and let $Y(\boldsymbol{x}) : \mathbb{R}^2 \to \mathbb{R}$ be the value of the luminance component at coordinate $\boldsymbol{x}$. We also use this notation for any other components. TV-regularization is for solving the following minimization problem.

$$Y_g = \underset{X}{\operatorname{argmin}} \sum |\nabla X(\boldsymbol{x})| + \frac{1}{2\lambda} \sum \{X(\boldsymbol{x}) - Y(\boldsymbol{x})\}^2 . \tag{4}$$

where Y is an original image and Y_g is a generated image. We call Y_g the geometry component, since Y_g only contains the geometry features (smooth regions and edges) of an original image. We call $Y_t = Y - Y_g$ the texture component because it presents the texture features included in an image. The solution for optimizing problem Eq. (4) was presented by Chambolle [14], who defines TV semi-norm $|\nabla \cdot|$ [14]. Note that the ranges of pixel values of geometry and texture components are $[0, 255]$ and $[-255, +255]$, respectively.

5. Correlation between Luminance and Chrominance

The color components of a natural image correlate with each other. Most chrominance changes are accompanied by luminance changes [7]–[12]. Therefore, each oscillation of the components in a local region has a similar shape and positive or negative correlation.

The natural image colorization [7] technique enhances the visual quality of a color image after the colorization process, assuming that the relationship between luminance and chrominance can be represented as a corresponding linear function. This correlation has also been utilized for color image coding [8], [9]. Compression techniques reduce the redundancies between components by predicting chrominance from luminance.

Spatio-chromatic correlation is used for color image coding [10]. $4 \times 4 \times 3$ bases are calculated by PCA or ICA using local color patches of $4 \times 4 \times 3$ [pixel]. Correlation between luminance and chrominance is suggested, since some upper bases contain chrominance change accompanied by luminance changes.

Also, Gershikov and Porat [11], [12] suggest the correlation between components of YC_bC_r color space and improve coding efficiency of JPEG/JPEG2000 by decorrelating. However, these methods need overheads to send the KLT matrix of each block or sub-band to the decoder.

Figures 1–3 show the local correlation between Y and C_b in separated geometry and texture components obtained by TV-regularization in Sect. 4. These lines mean the values extracted from the image "Food," in horizontal coordinate

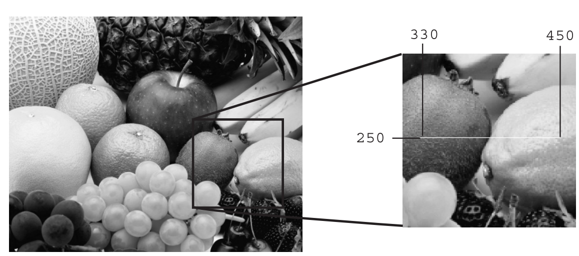


Fig. 1 Extracted area from image "Food" (white line).

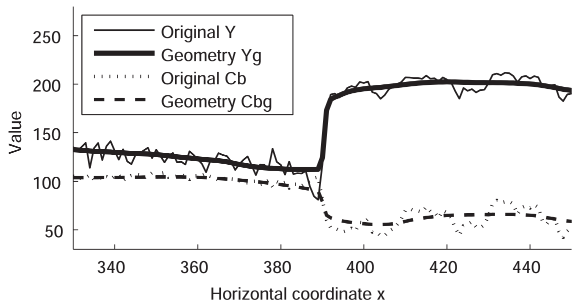


Fig. 2 Original image and its geometry components.

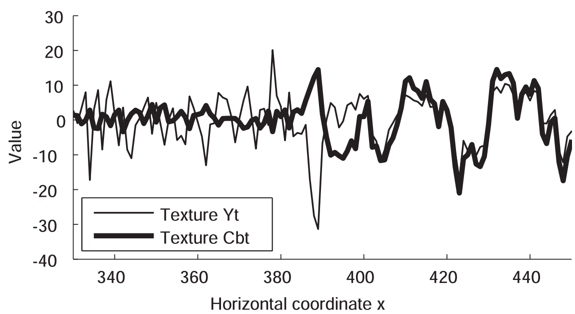


Fig. 3 Relationship between texture components of Y and C_b.

$x \in [330, 450]$ and fixed vertical coordinate $y = 250$. The values from $x < 390$ corresponding to Fig. 3 are the surface of a kiwi fruit, and the values $x \geq 390$ are the surface of a lemon. In Fig. 3, the kiwi area has a negative correlation between Y_t and C_{bt}. In contrast, the lemon area has a positive correlation. The key observation here is that these correlations change from negative to positive at the edge $x \simeq 390$, accompanied by a Y_g change (Fig. 2). This correlation is mostly caused by optical properties of the material surface and its magnitude is affected by gamma correction of YCbCr colorspace. However, even if we know these facts, it is difficult to decorrelate luminance and color components correctly.

6. Proposed Colorization-Based Coding

6.1 Overview of Proposed Method

Figure 4 shows an overview of the proposed colorization-based coding algorithm. The algorithm is described in Sects. 6.2–6.8. Here, we denote width and height of the input image as W [pixel] and H [pixel], respectively.

6.2 Geometry-Texture Separation

Our method separates each component of YC_bC_r into the geometry $Y_gC_{bg}C_{rg}$ and the texture $Y_tC_{bt}C_{rt}$ by Eq. (4). Then, the following equation is satisfied.

$$\begin{bmatrix} Y(\boldsymbol{x}) \\ C_b(\boldsymbol{x}) \\ C_r(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} Y_g(\boldsymbol{x}) \\ C_{bg}(\boldsymbol{x}) \\ C_{rg}(\boldsymbol{x}) \end{bmatrix} + \begin{bmatrix} Y_t(\boldsymbol{x}) \\ C_{bt}(\boldsymbol{x}) \\ C_{rt}(\boldsymbol{x}) \end{bmatrix}. \tag{5}$$

At the decoder side, only luminance Y is separated.

6.3 Correlation between Luminance and Chrominance in Texture Component

We denote the number of clusters as K, a coordinate in the image as $\boldsymbol{x} \in \mathbb{R}^2$, and an index of a segmented local region as $k(\boldsymbol{x}) : \mathbb{R}^2 \to \{1, \cdots, K\}$. By the observation in Sect. 5, if the clusters are properly segmented, we assume that the correlation in $Y_tC_{bt}C_{rt}$ can be formulated as

$$\begin{bmatrix} C_{bt}(\boldsymbol{x}) \\ C_{rt}(\boldsymbol{x}) \end{bmatrix} \simeq \begin{bmatrix} P_b^{k(\boldsymbol{x})} \\ P_r^{k(\boldsymbol{x})} \end{bmatrix} Y_t(\boldsymbol{x}). \tag{6}$$

where coefficients $P_b^{k(\boldsymbol{x})}$ and $P_r^{k(\boldsymbol{x})}$ represent the linear relationships between $Y_t(\boldsymbol{x})$ and $C_{bt}(\boldsymbol{x})$, or $C_{rt}(\boldsymbol{x})$, respectively. Note that $k(\boldsymbol{x})$ represents a superscript of coefficient $a_i^{k(\boldsymbol{x})}$, not an exponent.

Section 6.5 explains the clustering method, and Sect. 6.6 describes how to build arrays P_b^K and P_r^K that contain $P_b^{k(\boldsymbol{x})}$ and $P_r^{k(\boldsymbol{x})}$.

6.4 Extracting Vertex Information

The conventional method [3] needs to send the coordinates of two end points of each line segment, in addition to its chrominance values. In contrast, to improve the coding efficiency, our method uses point information that can only be obtained from the compressed luminance component Y_g'. Moreover, our method uses the vertices as color assignments, unlike the lines in the conventional method [3], since lines extracted from compressed luminance Y_g' could cut across the boundary of the areas that have different chrominance. On the encoder side, Y_g' can be obtained by local decoding. The algorithm of extracting vertex information from Y_g' is as follows.

1. Dividing compressed luminance Y_g' into $L \times L$ [pixel] non overlapped blocks, and putting vertex $\boldsymbol{x}$ at the center of each block. We denote the set of these vertices as V and the number of vertices as $|V| = HW/L^2$.
2. Quantized luminance value $Y_{gq}'(\boldsymbol{x}) \in \left\{0, 1, \cdots, q_Y^{max}\right\}$ is obtained by applying uniform quantization to $Y_g'(\boldsymbol{x})$. Where, $q_Y^{max} \in \mathbb{N}$ represents a maximum quantization level.
3. Making an array $I_V \in \mathbb{N}^{2\times|V|}$ by sorting the coordinates of the vertices $\forall \boldsymbol{x} \in V$ in ascending order of its corresponding quantized value $Y_{gq}'(\boldsymbol{x}) \in \left\{0, \cdots, q_Y^{max}\right\}$ (see Fig. 5).

Array I_V is the vertex information, as described above. The coordinates are stored in scanline order (first, ascending order of x coordinate, and second, ascending order of y) not in the whole area, but in each area that has the same level q_Y as described in processes 2 and 3. This ordering enables us to highly compress of cluster reference indexes (its details are described in Sect. 6.5) by run-length encoding. Arrows

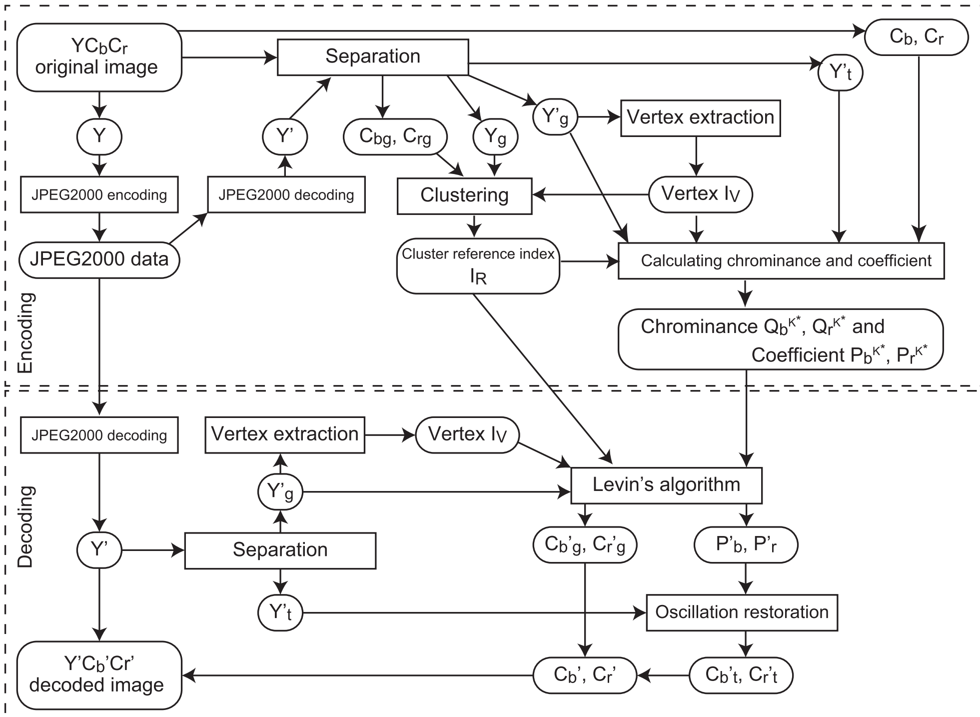


Fig. 4 Overview of proposed method.

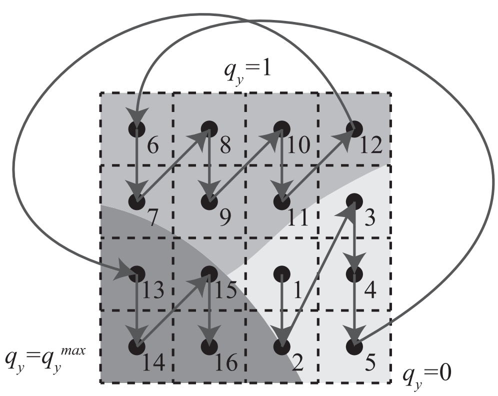


Fig. 5 Sort order of I_V.

in Fig. 5 show an example of order in the case of $|V| = 16$ and $q_Y^{max} = 2$. Vertex information I_V does not need to be sent to the decoder, since the same I_V can be obtained at the decoder side. This enables a significant rate reduction when compared with the conventional method [3].

6.5 Building Cluster Reference Index

At the encoder side, cluster reference indexes are built by using geometry components $Y_g C_{bg} C_{rg}$ of an original image YC_bC_r and vertex information I_V. The algorithm is as follows.

1. Normalizing vectors of $[Y_g(\boldsymbol{x}), C_{bg}(\boldsymbol{x}), C_{rg}(\boldsymbol{x})]$ by subtracting its average and dividing it by its standard deviation. Then, clustering the normalized vectors into K clusters ($k(\boldsymbol{x}) \in \{0, \cdots, K-1\}$) by using k-means method.
2. Storing cluster indexes $k(\boldsymbol{x})$ of all $\boldsymbol{x} \in V$ into an array $I_R \in \mathbb{N}^{|V|}$. Note that the orders of I_V and I_R are the same.

The array I_R represents the cluster reference indexes, as described above. I_R requires the following information volume to be stored.

$$V_R = |V| \lceil \log_2(K) \rceil \text{ [bit]} \tag{7}$$

Since segmented Y_g' and $Y_g C_{bg} C_{rg}$ have similar shapes, the same numbers occur in many consecutive elements in I_R. Therefore, the information volume of I_R can be compressed significantly.

I_R is transformed into

- L_R: Sequence whose element represents a run-length.
- C_R: Sequence whose element represents a cluster index.

by run-length encoding. Describing bits assigned to an element in L_R as b_R, an element in L_R ranges $\left\{1, \cdots, 2^{b_R}\right\}$, and that in C_R ranges $\{0, \cdots, K-1\}$. Since L_R and C_R are the functions of b_R, we can obtain the information volume V_R when we choose run-length coding with the best value of b_R by

$$V_R = \min_{b_R \in \{1,2,\ldots,b_R^{max}\}} \Big\{ \mathrm{len}(L_R(b_R)) b_R + \mathrm{len}(C_R(b_R)) \lceil \log_2(K) \rceil + \lceil \log_2(b_R^{max}) \rceil \Big\} \text{ [bit]}. \tag{8}$$

Where, $\mathrm{len}(\cdot)$ means length of a sequence.

For each input image, the proposed method compares

Eq. (7) with Eq. (8). Then, the method that can minimize V_R is adopted. An additional flag bit $f_R \in \{0, 1\}$ that states which method has been adopted is also sent to the decoder.

6.6 Calculating Chrominance and Coefficient Information

In this section, we describe how to calculate chrominance information $Q_b^K, Q_r^K \in \mathbb{R}^K$ and coefficient information $P_b^K, P_r^K \in \mathbb{R}^K$. Chrominance and coefficient information are arrays that have K elements of chrominance and coefficient corresponding to each cluster. Conventional colorization-based image coding developed by Ono et al. [6] presents how to calculate chrominance information that minimizes decoded error with fixed coordinates of assigned chrominance. However, since the chrominance component is the only information that is transmitted, Ono et al. did not investigate optimization for coefficient information. Moreover, they did not discuss clustering of color information. We extend the previous chrominance optimization method to clustered coefficients.

Let Ω ($|\Omega| = n$) be a set of all pixels and $V \subset \Omega$ be a set of pixels assigning chrominance. Focusing on C_b component, let $C_{bg}^{'} \in \mathbb{R}^n$ be a decoded geometry of C_b component, a vector $Q_b \in \mathbb{R}^{|V|}$ be chrominance information assigned at V, and $M \in \mathbb{R}^{n \times |V|}$ be a matrix whose column vectors are extracted from $A^{-1} \in \mathbb{R}^{n \times n}$ in Eq. (3) corresponding to the pixels in V.

For example, if $|V| = 2$ and chrominance value q_1 and q_2 are assigned to these two pixels. By using column vectors $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ corresponding to them, we can represent Eq. (3) as follows.

$$C_{bg}^{'} = \begin{bmatrix} \cdots & \boxed{\boldsymbol{a_1}} & \cdots & \boxed{\boldsymbol{a_2}} & \cdots \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ q_1 \\ \mathbf{0} \\ q_2 \\ \mathbf{0} \end{bmatrix} \tag{9}$$

$$= \begin{bmatrix} \boxed{\boldsymbol{a_1}} & \boxed{\boldsymbol{a_2}} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \tag{10}$$

$$= M Q_b. \tag{11}$$

Every vertex is clustered in our method. If q_1 and q_2 belong to the same cluster ($K = 1$), q_1 and q_2 must have the same chrominance ($q_1 = q_2$). Two column vectors $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are summed up into a new column vector $\boldsymbol{a_1} + \boldsymbol{a_2}$. Then, Eq. (11) is deformed as follows, using a new matrix $M_K \in \mathbb{R}^{n \times K}$ and a vector Q_b^K.

$$C_{bg}^{'} = \begin{bmatrix} \boxed{\boldsymbol{a_1} + \boldsymbol{a_2}} \end{bmatrix} \begin{bmatrix} q_1 \end{bmatrix} \tag{12}$$

$$= M_K Q_b^K. \tag{13}$$

If $|V| \geq 3$ or $K \geq 2$, M_K and Q_b^K are calculated in a similar way from Eq. (9) to Eq. (13).

Similarly, let $Y_t^{'} \in \mathbb{R}^n$ be a texture component of locally decoded luminance, $P_b^K \in \mathbb{R}^K$ be coefficient information, and $P_b^{'} \in \mathbb{R}^n$ be the vector propagated from P_b^K by M_K. A texture component of decoded chrominance $C_{bt}^{'} \in \mathbb{R}^n$ is formulated as follows.

$$\begin{aligned} C_{bt}^{'} &= P_b^{'} \odot Y_t^{'} && (14) \\ &= (M_K P_b^K) \odot Y_t^{'} && (15) \\ &= Y_t^{'} \odot (M_K P_b^K) && (16) \\ &= \mathrm{diag}(Y_t^{'}) M_K P_b^K. && (17) \end{aligned}$$

Here, $\odot$ means element-wise multiplication. Chrominance information Q_b^{K*} and coefficient information P_b^{K*} that minimize the square error between original chrominance C_b and decoded chrominance $(C_{bg}^{'} + C_{bt}^{'})$ is simultaneously obtained by solving the following minimization problem.

$$\begin{aligned} \{Q_b^{K*}, P_b^{K*}\} &= \underset{Q_b^K, P_b^K}{\mathrm{argmin}} \, \|C_b - (C_{bg}^{'} + C_{bt}^{'})\| \\ &= \underset{Q_b^K, P_b^K}{\mathrm{argmin}} \, \|C_b - (M_K Q_b^K + \mathrm{diag}(Y_t^{'}) M_K P_b^K)\| \end{aligned} \tag{18}$$

We can combine the second and third term of Eq. (18) by introducing matrix $N \in \mathbb{R}^{n \times 2K}$,

$$N = \begin{bmatrix} \boxed{M_K} & \boxed{\mathrm{diag}(Y_t^{'}) M_K} \end{bmatrix} \tag{19}$$

$$C_{bg}^{'} + C_{bt}^{'} = N \begin{bmatrix} Q_b^K \\ P_b^K \end{bmatrix} \tag{20}$$

Therefore, we can easily obtain the solution of Eq. (18) by calculating the Moore-Penrose pseudo inverse of N.

$$\begin{bmatrix} Q_b^{K*} \\ P_b^{K*} \end{bmatrix} = N^{\dagger} C_b \tag{21}$$

Where $N^{\dagger}$ denotes the Moore-Penrose pseudo inverse of N.

If we denote the information volume assigned to an element in Q_b^K and Q_r^K as b_Q [bit], total information volume of chrominance information can be expressed by $K \times 2b_Q$ [bit]. Similarly, if we let the information volume of the coefficient P_b^K and P_r^K be b_P [bit], total information volume of coefficient information becomes $K \times 2b_P$ [bit].

6.7 Sending Data

The following data are sent to the decoder side.

- Luminance component compressed by JPEG2000
- Header information
  - Number of iterations by TV-regularization: L_{TV}
  - Parameter of TV-regularization: λ_{TV}

  - Block size for extracting vertex: L
  - Maximum quantization level for extracting vertex: q_Y^{max}
  - Number of clusters: K
  - Flag bit of run-length encoding: f_R
  - Number of bits assigned to each element of run-length encoding: b_R
  - Number of bits assigned to each element of chrominance information: b_Q
  - Number of bits assigned to each element of coefficient information: b_P
- Cluster reference index: I_R (or L_R and V_R encoded from I_R)
- Chrominance information: Q_b^K and Q_r^K
- Coefficient information: P_b^K and P_r^K

6.8 Oscillation Restoration

By extracting vertex information I_V from the decoded luminance, and matching I_V with cluster reference index I_R, we can build the matrix M_K at the decoder side. Q_b^K and P_b^K are propagated by Eqs. (13) and (17), respectively. Using coefficient maps $P_b^{'}$ and $P_r^{'}$, and the texture component of decoded luminance $Y_t^{'}$, chrominance oscillation is restored by pixel-wise multiplication as follows.

$$\begin{bmatrix} C_{bt}^{'}(\boldsymbol{x}) \\ C_{rt}^{'}(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} P_b^{'}(\boldsymbol{x}) \\ P_r^{'}(\boldsymbol{x}) \end{bmatrix} Y_t^{'}(\boldsymbol{x}). \tag{22}$$

7. Simulation

7.1 Uncompressed Luminance

First, we compared several methods with flawless luminance component (without using any compression apart from lossless compression). Table 1 shows the encoder settings (Miyata et al. [3] and Inoue et al. [5] give the parameters). We used the 256×256 [pixel] color image "Mandrill" and evaluated the peak signal-to-noise ratio (PSNR) of C_b for several information volumes of the chrominance component. We compared our proposed method, conventional methods in [3], [5], JPEG, and JPEG2000. Note that JPEG and JPEG2000 encoding were performed for the chrominance component only.

Figure 6 shows that our proposed method coded images more efficiently than the conventional methods [3] and [5]. From this result, we can say that avoiding the coordinates information transmission is important to improve the coding efficiency. The proposed method also outperforms JPEG and JPEG2000, especially in terms of a high compression rate.

7.2 Compressed Luminance

The proposed method propagates chrominance values and coefficients on a luminance component by using a colorization method [1]. If the luminance is highly compressed, the degradation of luminance may affect the decoding process of the chrominance. For example, if we use the compressed luminance component that loses sharp edges, chrominance may be propagated beyond the borders between the appropriate and inappropriate areas. Moreover, if we use the luminance component that loses oscillation like the texture component, chrominance fails to restore their oscillations.

In this section, we compare (a) JPEG2000 with (b) the proposed method using JPEG2000 to compress the luminance component. Their performances are evaluated using the following processes. Let PSNR(X_1X_2) be a PSNR calculated from average of two mean squared errors (MSE) — between X_1 and its corresponding original component, and between X_2 and its original.

Similarly, we denote an average of structured similarity (SSIM) [15] between X_1 and its original and SSIM between X_2 and its original as SSIM(X_1X_2). The SSIM is a well-known and well-used quality metric used to measure the

Table 1 Encoder settings.

Method	Parameter	Value
Proposed	L	8
	K	5, 10, 15, 20, 25, 30, 35, 40
	b_P	8
	L_{TV}	100
	λ_{TV}	0.2
	q_Y^{max}	8
	b_Q	8
	b_R^{max}	8
Conventional [3]	V	15
	th	4, 6, 8, 12, 16, 20, 24, 28
	S	0.25
Conventional [5]	V	15
	th	4, 6, 8, 12, 14, 16, 20
	S	0.25
	b_P	8
JPEG	q	$1, 3, \cdots, 27$
JPEG2000	rate	0.030, 0.031, 0.032, 0.033, 0.034, 0.035, 0.036, 0.038, 0.040, 0.045, 0.050, 0.055, 0.06, 0.07, 0.08, 0.1

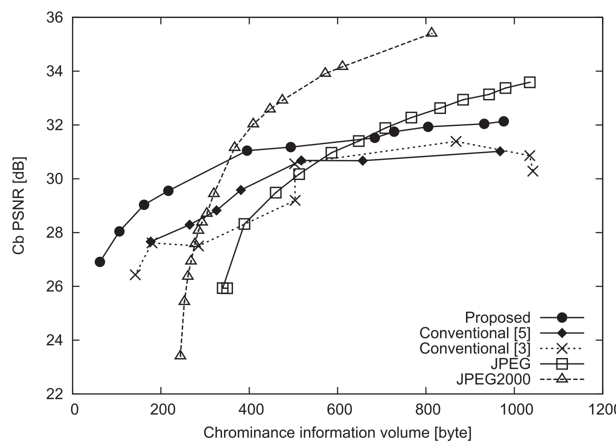
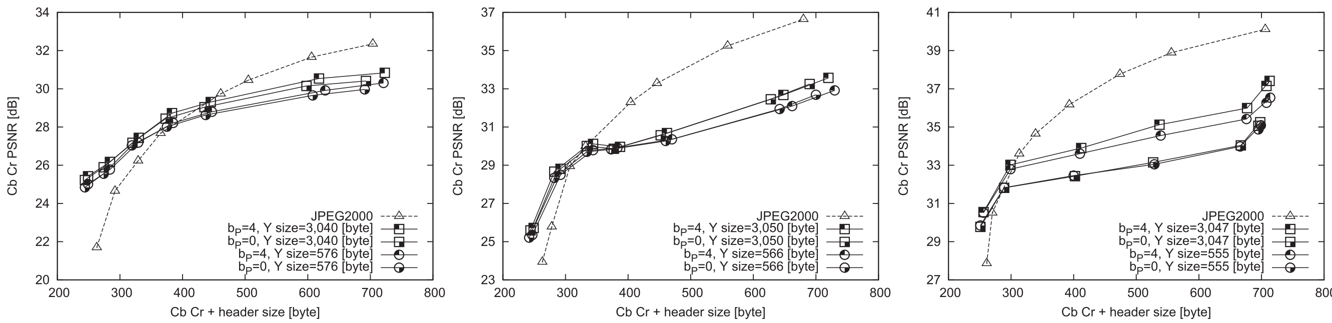
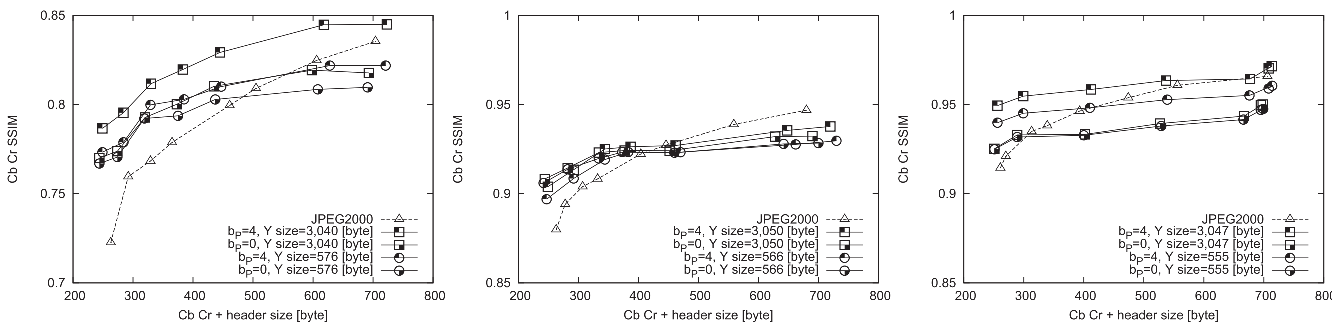


Fig. 6 Comparison of coding efficiency (uncompressed Y, "Mandrill").



(a) "Mandrill" (b) "Parrots" (c) "Sunset"

Fig. 7 Comparison of CbCr PSNR (compressed Y).



(a) "Mandrill" (b) "Parrots" (c) "Sunset"

Fig. 8 Comparison of CbCr SSIM (compressed Y).

Table 2 Encoder settings.

Parameter	Value
(L, K)	(48, 5), (32, 10), (16, 10), (12, 10), (8, 10), (8, 20), (8, 30)
b_P	0, 4
L_{TV}	100
λ_{TV}	0.2
q_Y^{max}	8
b_Q	8
b_R^{max}	8
$r_{\tilde{Y}}$	0.1, 0.4

similarity between two images because it is a better model of the human quality perception than PSNR.

1. Preparing several compressed images $\hat{Y}\hat{C}_b\hat{C}_r$ using JPEG2000, by changing luminance bitrate and chrominance bitrate (sum of C_bC_r).
2. Obtaining $\tilde{Y}$ by JPEG2000 compression, with compression rate $r_{\tilde{Y}}$, applied for the luminance component Y only. Setting the compression rate to satisfy PSNR($\hat{Y}$) = PSNR($\tilde{Y}$). Compressing the chrominance component by our proposed method with the compressed luminance component $\tilde{Y}$. Then, obtaining all decoded components $\tilde{Y}C_b^{'}C_r^{'}$.
3. Evaluating PSNR($\hat{C}_b\hat{C}_r$), PSNR($C_b^{'}C_r^{'}$), SSIM($\hat{C}_b\hat{C}_r$), and SSIM($C_b^{'}C_r^{'}$).

Table 2 shows the encoder settings.

Results are shown in Figs. 7 and 8. The horizontal axis represents sum information volume of the chrominance components and the header. For JPEG2000, this information amount is the sum of (a-1) and (a-3), and for the proposed method, it is the sum of (b-1), (b-3), and (b-4) (see Fig. 9). The vertical axis represents the quality of the decoded images of the proposed method (PSNR($C_b^{'}C_r^{'}$) and SSIM($C_b^{'}C_r^{'}$)) and JPEG2000 (PSNR($\hat{C}_b\hat{C}_r$) and SSIM($\hat{C}_b\hat{C}_r$)). In many images, our method coded more efficiently than JPEG2000, especially in the case of the combination of high bit rate luminance and low bit rate chrominance. Figure 7 shows that, in "Mandrill" and "Parrots" images, we obtained +3 [dB] PSNR gain by using our proposed method. Moreover, Fig. 8 shows that there is a remarkable difference between $b_P = 0$ and $b_P = 4$ on SSIM metrics, and which means that our main idea can improve the perceptual quality drastically. Output images are shown

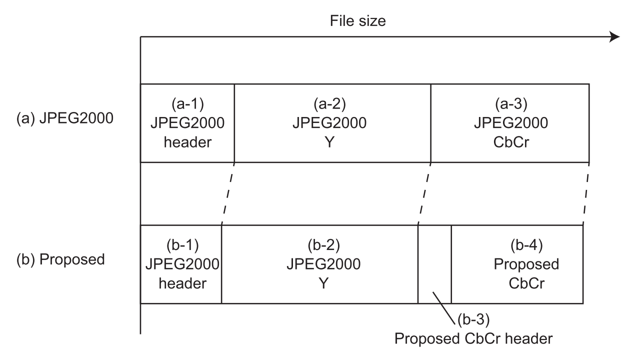
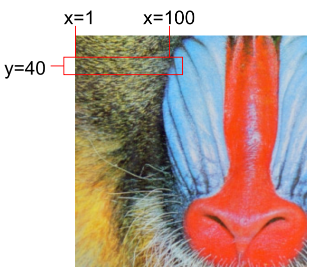
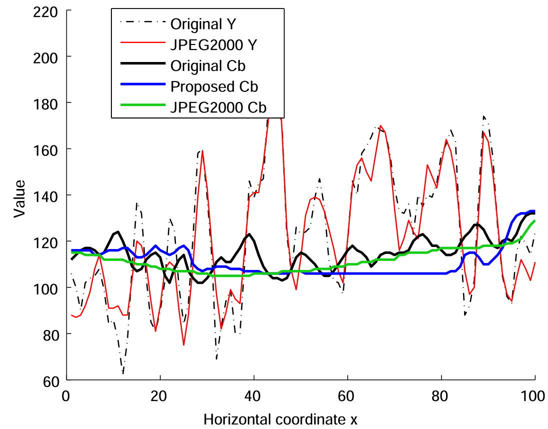


Fig. 9 Comparison of information volume.

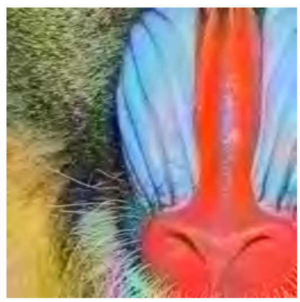


(a) Extracted area from image "Mandrill", horizontal coordinate $x \in [1, 100]$, vertical coordinate $y = 40$.



(b) Local oscillation restoration by proposed method (horizontal axis: horizontal coordinate x, vertical axis: pixel value).

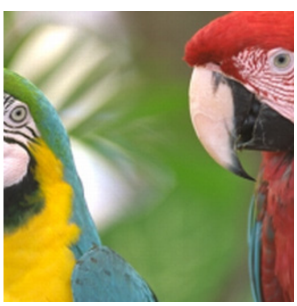
(c) Original.



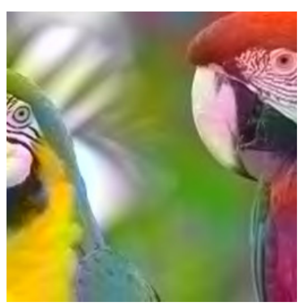
(d) JPEG2000 (C_bC_r: 331 [byte], 26.19 [dB], 0.7679).

(e) Proposed (C_bC_r: 330 [byte], 27.43 [dB], 0.8117).

Fig. 10 Mandrill ($r_{\tilde{Y}} = 0.4$, Y: 3,040 [byte], 28.62 [dB]).



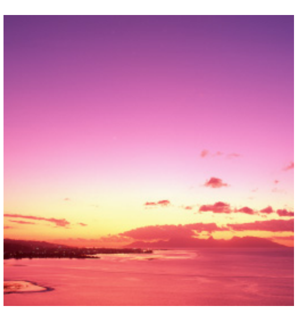
(a) Original.



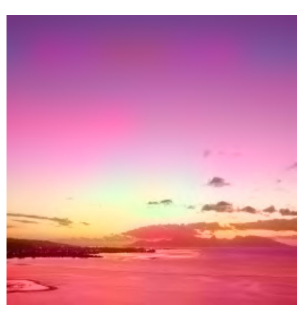
(b) JPEG2000 (C_bC_r: 287 [byte], 26.85 [dB], 0.8923).

(c) Proposed (C_bC_r: 292 [byte], 28.81 [dB], 0.9135).

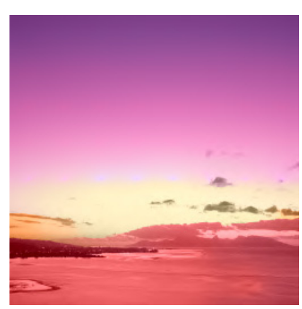
Fig. 11 Parrots ($r_{\tilde{Y}} = 0.4$, Y: 3,050 [byte], 38.41 [dB]).



(a) Original.



(b) JPEG2000 (C_bC_r: 306 [byte], 32.63 [dB], 0.9277).



(c) Proposed (C_bC_r: 299 [byte], 33.02 [dB], 0.9547).

Fig. 12 Sunset ($r_{\tilde{Y}} = 0.4$, Y: 3,047 [byte], 44.33 [dB]).

in Fig. 10 (c) through (e), Figs. 11 and 12. JPEG2000 generated different colors from those of neighboring areas, for example, on left side of the nose in "Mandrill," in the background of "Parrots," and in the sky of "Sunset." In contrast, our method restored natural colors and gradations and did not generate such noises.

Figure 10 (b) shows the characteristic of the local oscillation restoration of the proposed method. These chrominance values are extracted from Fig. 10 (c) through Fig. 10 (e), and the location of these pixels is shown in Fig. 10 (a). JPEG2000 is more smoothed than the original values. On the other hand, the chrominance of our proposed method contains the local oscillation and has a similar waveform to the original, especially in $x \in [1, 40]$.

8. Conclusion

In this paper, we presented a novel colorization-based image coding method based on the correlation between luminance and chrominance in separated texture components. Since we focused on the relationship between edges of the geometry component and the boundary of positive correlation and negative correlation, our proposed method efficiently compressed coefficients that represent the correlation. Furthermore, our proposed method enables drastic bit-rate reduction by sharing same vertices at both the encoder and decoder sides. The experimental results showed that our coding method outperforms the conventional method, JPEG, and JPEG2000, especially when high bit rate luminance and low bit rate chrominance are combined.

References

[1] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," ACM Trans. Graphics, vol.23, pp.689–694, 2004.

[2] T. Takahama, T. Horiuchi, and H. Kotera "Improvement on colorization accuracy by partitioning algorithm in CIELAB color space," PCM 2004, LNCS 3332, pp.794–801, 2004.

[3] T. Miyata, Y. Komiyama, Y. Sakai, and Y. Inazumi, "Novel inverse colorization for image compression," Picture Coding Symposium 2009, pp.101–104, 2009.

[4] M. Nishi, T. Horiuchi, and H. Kotera, "A novel picture coding using colorization technique," IS & T's International Conference on Digital Printing Technologies (NIP21), pp.380–383, 2005.

[5] Y. Inoue, T. Miyata, and Y. Sakai, "Improving colorization-based coding by using local correlation between luminance and chrominance in texture component," Picture Coding Symposium 2010, pp.398–401, 2010.

[6] S. Ono, T. Miyata, and Y. Sakai, "Colorization-based coding by focusing on characteristics of colorization bases," Picture Coding Symposium 2010, pp.230–233, 2010.

[7] L. Qing, F. Wen, D. Cohen-Or, L. Liang, Y.Q. Xu, and H. Shum, "Natural image colorization," in Rendering Techniques 2007 (Proc. Eurographics Symp. on Rendering), pp.309–320, 2007.

[8] J.O. Limb and C.B. Rubinstein, "Plateau coding of the chrominance component of color picture signals," IEEE Trans. Commun., vol.22, no.6, pp.812–820, 1974.

[9] X. San, H. Cai, and J. Li, "Color image coding by using inter-color correlation," IEEE International Conference on Image Processing, pp.3117–3120, 2006.

[10] M.S. Drew and S. Bergner, "Spatio-chrominance decorrelation for color image compression," Technical Report, School of Computing Science, Simon Fraser Univiersity, 2007.

[11] E. Gershikov and M. Porat, "Does decorrelation really improve color image compression?," International Conference on Systems Theory and Scientific Computation (ISTASC05), 2005.

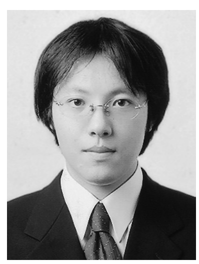
[12] E. Gershikov and M. Porat, "Optimal color image compression using localized color components transforms," European Signal Processing Conference, in CD-ROM, 2008.

[13] L.I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D, vol.60, no.1-4, pp.259–268, 1992.

[14] A. Chambolle, "An algorithm for total variation minimization and applications," J. Mathematical Imaging and Vision, vol.20, no.1-2, pp.89–97, 2004.

[15] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol.13, no.4, pp.600–612, 2004.

Yoshitaka Inoue received a B.E. degree from Waseda University in 2009. He joined the Dept. of Communications and Integrated Systems at the Tokyo Institute of Technology in 2009 where is currently undertaking a master's degree. His research interests include image processing and image coding.



Takamichi Miyata received B.E. and M.E. degrees from the University of Toyama in 2001 and 2003, respectively, and a Ph.D. degree from the Tokyo Institute of Technology in 2006. In 2006, he joined the Dept. of Communications and Integrated Systems at Tokyo Institute of Technology as an assistant professor. His current research interests include image processing and image coding.

Yoshinori Sakai received B.Eng., and Dr.Eng. in electrical engineering from the University of Tokyo in 1969 and 1974 respectively. From 1974 to 1987, he was employed by Nippon Telegraph and Telephone Public Corporation (NTT). From 1987 to 1990, he was an Associate Professor at the Tokyo Institute of Technology, where he is a Professor. He designed many communication systems, for example digital transmission system over existing FDM network, facsimile communication network and teleconference system at NTT lab. His current research interests include image information retrieval, network streaming and content distribution over the Internet. He received the best author award from the ITV Japan, the best paper award from IIEEJ Japan and the achievement award from IEICE. He is now Director, Publications of IEICE.