ABSTRACT

Recently, a novel approach to color image compression based on colorization has been presented. Although the conventional method of colorization-based coding outperforms JPEG in terms of subjective quality, the decoded chrominance components lose the local oscillation that the original images had. A large number of color assignments is required to restore these oscillations.

We focus on the local correlation that exists between luminance and chrominance in separated texture components, and we present a new colorization-based coding method. Experimental results showed that our coding method can restore the oscillation and improve the coding efficiency compared with the conventional method.

Index Terms— image coding, colorization, total variation, correlation between luminance and chrominance

1. INTRODUCTION

Colorization is a process that adds color to a grayscale image [1]. Recently, a novel approach to color image compression based on colorization has been presented. The conventional method [2] segments clusters of chrominance components of an input color image and assigns colors to the clusters at the encoder. The chrominance components are propagated from the color assignation by a colorization technique at the decoder. Although the conventional method of colorization-based coding outperforms JPEG in terms of visual quality, the decoded chrominance components lose the local oscillation that the original images had. A large number of color assignments is required to restore these oscillations precisely. Previous studies [3, 4] suggest that luminance and chrominance components of a natural image correlate, and most chrominance changes are accompanied by luminance changes.

In this paper, we focus on the correlation that exists between luminance and chrominance at a local area in an image. We assume that the correlation exists in texture components separated by using total variation regularization [5, 6], and we present a new colorization-based coding method that restores oscillations. The key idea of our method is to apply different compression techniques to the geometry and texture components. The texture components are compressed into coefficients that represent the correlation between luminance and chrominance, and the geometry components are compressed by the conventional method.

2. COLORIZATION-BASED CODING

2.1. Encoding / Decoding

Figure 1 shows an overview of the conventional colorization-based coding algorithm [2]. The encoding algorithm is as follows.

1. Original chrominance (Cb and Cr) components are rescaled by factor $S \ (S < 1)$.

2. $C_b$-$C_r$ clustering: Original chrominance components are segmented into clusters by using the farthest-point algorithm, minimizing the maximum radius (distance to the cluster center) in $C_b$-$C_r$ space. The distance between each cluster representative color and the color of any pixel in that cluster is less than given parameter $th$.

3. Line extraction: Lines are added to color assignation data repeatedly until the distance between each corresponding pixel of the decoded chrominance and the original is less than $th$. In this case, only lines are
added that can decrease the square error of the chrominance components more than the given parameter $V$. More detailed information of this step is described in Ref. [2].

In the conventional method, the compression data consist of the (a) color pallet, (b) color assignment, and (c) luminance component. The color pallet has each cluster representative color. The color assignment contains two coordinates that represent a line segment (coordinates of a start point and an end point) and an index number referring to the color pallet. To compress the luminance component, one can use any standardized compression method, such as JPEG/JPEG2000. The decoding algorithm is as follows.

1. Luminance component $Y'$ is decoded.
2. Color line segments are drawn on $Y'$ by using the color assignment and the color pallet.
3. Levin’s algorithm: Chrominance components are propagated from color line segments on $Y'$.

Here, $Y'$, $C_{bt}'$, and $C_{tr}'$ mean the decoded components of $Y$, $C_b$, and $C_r$, respectively.

2.2. Problem

The colorization method by Levin et al. [1] is used to decode chrominance components based on the assumption that neighboring pixels that have similar luminance have similar chrominance. However, the decoded chrominance components lose the local oscillation that the original images had. Since the local oscillations represent the textures and shadows of materials, the loss of such oscillations leads to considerable degradation of subjective quality. A large number of color assignments is required to restore these oscillations precisely. However, the increase of the number of color assignments makes the compressed data size large.

3. PROPOSED COLORIZATION-BASED CODING

3.1. Geometry-texture separation

Denoising methods that use total variation regularization have been proposed [5, 6]. Let $x \in \mathbb{R}^2$ be a coordinate of the image, and let $Y(x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ be the value of the luminance component at coordinate $x$. We also use this notation for any other components. Total variation regularization is for solving the following minimization problem.

$$Y_g = \arg \min_X \sum_x |\nabla X(x)| + \frac{1}{2\lambda} \sum_x \{X(x) - Y(x)\}^2.$$  (1)

Here, $Y$ is an original image and $Y_g$ is a generated image. We call $Y_g$ the geometry component, since one can recognize the geometry of an image, such as regions and edges, in $Y_g$. We call $Y_t = Y - Y_g$ the texture component because it presents the texture of each material included in an image.

In this paper, each component of $L = \{Y, C_b, C_r\}$ is separated into geometry $L_g = \{Y_g, C_{bg}, C_{rg}\}$ and texture $L_t = \{Y_t, C_{bt}, C_{rt}\}$. The ranges of the geometry and texture are $[0, 255]$ and $[-255, +255]$, respectively. Note that we have the following relationship between the geometry and texture portions of each component.

$$L(x) = L_g(x) + L_t(x).$$  (2)

3.2. Correlation between luminance and chrominance

The chrominance changes of a natural image correlate. Most chrominance changes are accompanied by luminance changes [3, 4]. 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 [3, 4]. Compression techniques reduce the redundancies between components by predicting chrominance from luminance.

Figure 2 shows the local correlation between $Y$ and $C_b$ in separated geometry and texture components (these lines mean the values extracted from image “Food,” horizontal coordinate $x \in [330, 450]$, and fixed vertical coordinate $y = 250$). The values from $x < 60$ corresponding to Fig. 2 are the surface of a kiwi fruit, and the values $x \geq 60$ are the surface of a lemon. In Fig. 2 (b), 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 $x \approx 60$, accompanied by a $Y_g$ change (Fig. 2 (a)). This correlation is mostly caused by optical properties of the material surface and gamma correction at RGB colorspace. However, even if we know these facts, it is difficult to decorrelate luminance and color components correctly.

Describing the number of clusters as $K$ and an index of a segmented local region as $k(x) : \mathbb{R}^2 \rightarrow \{1, \cdots, K\}$, we assume that the correlation in $Y_t C_{bt} C_{rt}$ can be formulated as

$$\begin{bmatrix} C_{bt}(x) \\ C_{rt}(x) \end{bmatrix} = \begin{bmatrix} a_1^{k(x)} \\ a_2^{k(x)} \end{bmatrix} Y_t(x).$$  (3)

Here, $k(x)$ is, for example, given by applying the same clustering method as the encoding step 2 in 2.1 to $C_{bt} C_{rt}$, and the number of clusters $K$ depends on the given $th$, $a_1^{k(x)}$ is calculated by the least mean square method in each local area $k(x)$ and represents the linear relationship between $Y_t(x)$ and $C_{bt}(x) (C_{rt}(x))$. Note that $k(x)$ represents a superscript of coefficient $a_1^{k(x)}$, not an exponent.
3.3. Encoding / Decoding

We present a novel colorization-based coding that use geometry-texture separation and the correlation between luminance and chrominance. The key idea of our method is to apply different compression techniques to the geometry and texture components of an original input image. The geometry component can be efficiently compressed by the conventional method since it has no oscillation and needs less color assignation than the case of applying the conventional method to the original image. The texture component is compressed into coefficients that represent the correlation (see Eq. 4).

Figure 3 shows an overview of the proposed colorization-based coding algorithm. The encoding algorithm is as follows.

1. Separation: $Y C_b C_r$ is separated into geometry $Y g C_{bg} C_{rg}$ and texture $Y t C_{bg} C_{rt}$ components by using total variation minimization [6].

2. $C_{bg} C_{rg}$ is compressed by the conventional method [2], provided that the representative colors are chosen in $C_b C_r$.

3. Coefficient extraction: Coefficients $a_k(x)$ in the $k$-th local region are calculated by the least mean square method. The regions in this step are the clusters produced in the previous step.

The decoding algorithm is as follows.

1. Separation: $Y' g$ is decoded and separated into the geometry $Y_g$ and texture $Y_t$.

2. Line segments that have chrominance and coefficients are drawn on $Y_g$.

3. Levin’s algorithm: Chrominance components $C_{bg} C_{rg}$ are propagated from line segments on $Y_g$. A coefficient map $A_i(x) : \mathbb{R}^2 \rightarrow \mathbb{R}(i \in \{1, 2\})$ is also obtained by propagating $a_k(x)$ on $Y'_g$.

4. Oscillation restoration: $Y t C_{bg} C_{rt}$ is restored by the following multiplication at each coordinate:

$$
\begin{bmatrix}
    C_{bg}'(x) \\
    C_{rt}'(x)
\end{bmatrix} =
\begin{bmatrix}
    A_1(x) \\
    A_2(x)
\end{bmatrix}
Y_t(x).
$$

5. $C_b C_r$ is restored from $C_{bg}' C_{rg}'$ and $C_{bg}' C_{rt}'$ (see Eq. 2).

4. EXPERIMENTAL RESULTS

4.1. Settings

We compared the proposed method with the conventional colorization coding [2] and JPEG. Four natural color images were used as original images. In this experiment, we assumed that luminance component $Y$ is sent to the decoder side without any compression (i.e., $Y' = Y$). We set a threshold of contribution $V = 30$, a scaling factor $S = 1/4$, a minimum color error $th \in \{4, 6, 8, 12, 16, 20, 24\}$, and $\lambda = 0.2$
Proposed
Conventional
JPEG

0 1000 2000
24
26
28
30
32
Chrominance information volume [bytes]
PSNR [dB]

(a) Mountain
(b) Mandrill

(c) Food
(d) Night

Fig. 4. Comparison of coding efficiency.

Fig. 5. Local oscillation restoration by proposed method (horizontal axis: horizontal coordinate \(x\), vertical axis: pixel value).

4.2. Results

Figure 4 shows the experimental results of each image. The horizontal axis represents the information volume of the chrominance, and the vertical one represents PSNR of \(C_b\). From these graphs, we can see that the proposed method outperforms the conventional method in most images, especially in (a). The proposed method also demonstrates higher coding efficiency, in comparison with JPEG at a high compression rate. Figure 5 shows local oscillations (\(C_b\) component of image “Mountain,” horizontal coordinate \(x\) ∈ [380, 450] and vertical coordinate \(y = 380\)). The conventional method in Fig. 5 has 870 [byte] and 32.17 [dB] of PSNR, and the proposed method has 615 [byte] and 33.90 [dB]. Although the conventional method radically smoothed the local oscillation, the proposed method can represent the oscillations that the original image has.

5. CONCLUSION

In this paper, we presented a novel colorization-based coding method based on the correlation between luminance and chrominance in separated texture components. Since we fo-