Colorization-Based Compression Using Optimization

Sukho Lee, Sang-Wook Park, Paul Oh, and Moon Gi Kang, *Member, IEEE*

***Abstract*— In this paper, we formulate the colorization-based coding problem into an optimization problem, i.e., an L_1 minimization problem. In colorization-based coding, the encoder chooses a few representative pixels (RP) for which the chrominance values and the positions are sent to the decoder, whereas in the decoder, the chrominance values for all the pixels are reconstructed by colorization methods. The main issue in colorization-based coding is how to extract the RP well therefore the compression rate and the quality of the reconstructed color image becomes good. By formulating the colorization-based coding into an L_1 minimization problem, it is guaranteed that, given the colorization matrix, the chosen set of RP becomes the optimal set in the sense that it minimizes the error between the original and the reconstructed color image. In other words, for a fixed error value and a given colorization matrix, the chosen set of RP is the smallest set possible. We also propose a method to construct the colorization matrix that colorizes the image in a multiscale manner. This, combined with the proposed RP extraction method, allows us to choose a very small set of RP. It is shown experimentally that the proposed method outperforms conventional colorization-based coding methods as well as the JPEG standard and is comparable with the JPEG2000 compression standard, both in terms of the compression rate and the quality of the reconstructed color image.**

***Index Terms*— Colorization, compressive sensing, image compression, image filtering, reconstruction.**

I. INTRODUCTION

RECENTLY, a new compression technique for color images, which is based on the use of colorization methods, has been proposed [1]–[4]. Previously, several colorization methods [5] have been proposed to colorize grayscale images using only a few representative pixels provided by the user. The main task in colorization based compression is to automatically extract these few representative pixels in the encoder. In other words, the encoder selects the pixels required for the colorization process, which are called representative pixels (RP) in [4], and maintains the color information only for these RP. The position vectors and the chrominance values are sent to the decoder only for the RP set together with the luminance channel, which is compressed by conventional compression techniques. Then, the decoder restores the color information for the remaining pixels using colorization methods.

The main issue in colorization based coding is how to extract the RP set so that the compression rate and the quality of the restored color image becomes good. Several methods have been proposed to this end [1]–[4]. All these methods take an iterative approach. In these methods, first, a random set of RP is selected. Then, a tentative color image is reconstructed using the RP set, and the quality of the reconstructed color image is evaluated by comparing it with the original color image. Additive RP are extracted from regions where the quality does not satisfy a certain criterion using RP extraction methods, while redundant RP are reduced using RP reduction methods. However, the set of RP may still contain redundant pixels or some required pixels may be missing.

The main contribution of this paper is that we formulated the RP selection problem into an optimization problem, that is, an L_1 minimization problem. The selection of the RP is optimal with respect to the given colorization matrix in the sense that the difference error between the original color image and the reconstructed color image becomes minimum with respect to the L_2 norm error. Furthermore, the number of pixels in the RP set is also minimized by the L_1 minimization. The optimal set of RP is obtained by a single minimization step, and does not require any refinement, i.e., any additional RP extraction/reduction methods. Therefore, there is no need for iteration. Furthermore, there is no need to use a geometric method such as defining line segments or squares as in [1]–[4].

The optimization problem can also be considered as a variational approach, and therefore, the rich research results of the variational approach in image processing can be used in the colorization based coding problem.

We also propose a construction method of the colorization matrix, which, combined with the proposed RP extraction method, produces a high quality reconstructed color image. It will be shown experimentally that the proposed scheme compresses the color image with higher compression rate than the conventional JPEG standard as well as other colorization based coding methods, and is comparable to the JPEG2000 standard even without using complex entropy coding for the proposed method.

II. RELATED WORKS

To understand the proposed method, three major related works have to be explained.

Manuscript received September 22, 2012; revised February 22, 2013; accepted March 3, 2013. Date of publication March 20, 2013; date of current version May 13, 2013. This work was supported by the Mid-Career Researcher Program under an NRF Grant 2012R1A2A4A01003732 funded by the MEST. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kenneth Lam.

S. Lee is with the Department of Software Engineering, Dongseo University, Busan 617-716, Korea (e-mail: petrasuk@gmail.com).

S.-W. Park, P. Oh, and M. G. Kang are with the Department of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Korea (e-mail: scrambls@hanmail.net; phoenix820@naver.com; mkang@yonsei.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2253486

A. Levin's Colorization Technique

In [5], Levin *et al*'s propose a colorization algorithm, which reconstructs the colors in the decoder using the color information for only a few representative pixels (RP) and the gray image which contains the luminance information. For example, using the YCbCr color space, the colorization problem reconstructs all the Cb and Cr components, given the Y luminance component and the Cb and Cr information for a few RP. Following the notation in [4], we denoted $\mathbf{y}$ as the luminance vector, $\mathbf{u}$ as the solution vector, i.e., the vector containing the color components to be reconstructed in the decoder, and $\mathbf{x}$ as the vector which contains the color values only at the positions of the RP, and zeros at the other positions. The vectors $\mathbf{y}$, $\mathbf{u}$, and $\mathbf{x}$ are all in raster-scan order. The cost function defined by Levin *et al.* is

$$J(\mathbf{u}) = \|\mathbf{x} - A\mathbf{u}\|^2 \tag{1}$$

which has to be minimized with respect to $\mathbf{u}$. Here, $A = I - W$, where I is an $n \times n$ identity matrix, n is the number of pixels in $\mathbf{u}$, and W is an $n \times n$ matrix containing the w'_{rs} weighting components. The w'_{rs} weighting components are

$$w'_{rs} = \begin{cases} 0 & \text{if } r \in \Omega \\ w_{rs} & \text{otherwise} \end{cases}$$

where

$$w_{rs} \propto e^{(y(r)-y(s))^2/2\sigma_r^2}. \tag{2}$$

Here, Ω denotes the set of the positions of the RP, σ_r^2 is a small positive value, and w_{rs} is the weighting component between the pixels at the r's and the s's positions, where $s \in N(r)$, and $N(r)$ is the 8-neighborhood of the r's pixel. Furthermore, $y(r)$ and $y(s)$ are the luminance values at the r's and the s's positions in the luminance vector $\mathbf{y}$, respectively. The minimizer of (1) can be explicitly computed as

$$\mathbf{u} = A^{-1}\mathbf{x}. \tag{3}$$

B. Colorization-Based Compression Techniques

As mentioned in the introduction, the main function of colorization based coding is the extraction of the RP. Previous colorization based coding methods use an iterative approach to extract the RP. In these approaches, first, an *a priori* temporary set of RP is usually selected. This *a priori* selection is manual and causes a redundant or insufficient set of RP. Therefore, redundant RP have to be eliminated, and required RP have to be additionally extracted by additional RP elimination/extraction methods.

In [1] and [2], new pixels are added to the initial set of RP by iterative selection based on machine learning, while in [3], the RP is selected iteratively constrained to a set of color line segments. In [4], redundant RP are reduced and required RP are extracted iteratively based on the characteristics of the colorization basis. However, after using these additional RP extraction/reduction methods, it is still not guaranteed that the resulting set of RP is optimal. After each step of RP extraction, the A matrix is constructed, and the A^{-1} matrix is obtained by taking the inverse of A. Then, a tentative color image is reconstructed by (3) which is then compared with the original color image to decide if more RP have to be extracted. The matrix A has to be constructed after each step of RP extraction, since A differs for different sets of RP.

C. L_1 Minimization Model

In many applications, it is necessary to obtain a highly sparse signal $\mathbf{x}$, i.e., a signal with very few nonzero components, which produces the measurement $\mathbf{b}$, given a certain system (matrix) A: $A\mathbf{x} = \mathbf{b}$. This problem can be formulated as an L_0 minimization problem:

$$\underset{\mathbf{x}}{\operatorname{argmin}} |\mathbf{x}|_0, \text{ s.t. } \mathbf{b} = \mathrm{A}\mathbf{x} \tag{4}$$

where the $|\mathbf{x}|_0$ norm measures the number of nonzero components in $\mathbf{x}$. However, this problem is very difficult to solve, since it is a combinatorial optimization problem with prohibitive complexity. Recently, it has been established theoretically that the solution $\mathbf{x}$ can be obtained also as a solution of an L_1 minimization problem if A satisfies the RIP (Restricted Isotropy Property) condition [6]–[11]

$$\underset{\mathbf{x}}{\operatorname{argmin}} |\mathbf{x}|_1, \text{ s.t. } \mathbf{b} = \mathrm{A}\mathbf{x}. \tag{5}$$

The L_1 minimization problem can be solved easily by tractable linear programming. The L_1 minimization was popularized by the work in [6] and is now widely used in the image processing area, especially in the total variation minimization and the compressive sensing area. Stability results have been established for the L_1 minimization model in [8]–[11]. One of the major contributions of this paper is that we formulated the RP selection problem into an L_1 minimization problem.

III. Proposed Method

While most colorization based coding methods try to extract the RP set by using an iterative approach, we formulate the RP selection problem into an L_1 minimization problem. An essential prerequisite for this is that the colorization matrix has to be determined beforehand. We will first explain why the L_1 minimization problem suits the RP selection problem well. Then, we propose a method to determine the colorization matrix from the given luminance channel before the RP selection.

A. Overall System Diagram

Figure 1 shows the overall system diagram of the proposed method. The details of the components are described in the following subsections. In the encoder, the original color image is first decomposed into its luminance channel and its chrominance channels. The luminance channel is compressed using conventional one-channel compression techniques, e.g., JPEG standard, and its discrete Fourier or Wavelet coefficients are sent to the decoder. Then, in the encoder, the colorization matrix C is constructed by performing a multi-scale meanshift segmentation on the decompressed luminance channel. The decompressed luminance channel is used to consists with that in the decoder. Using this matrix C and the original chrominance values obtained from the original color image, the RP

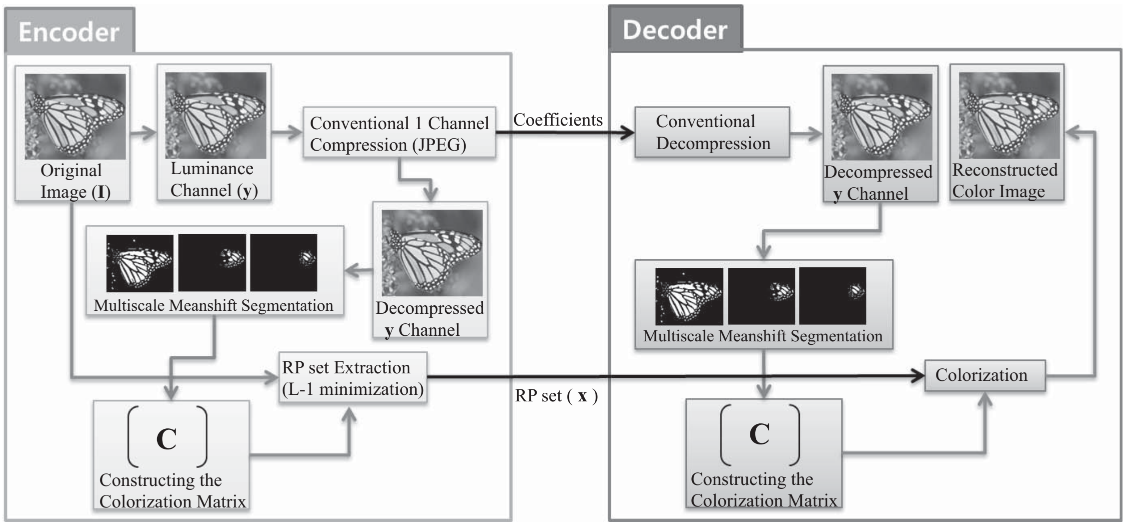


Fig. 1. Overall system diagram from encoder to decoder of the proposed compression framework.

set is extracted by solving an optimization problem, i.e., an L_1 minimization problem. This RP set is sent to the decoder, where the colorization matrix C is also reconstructed from the decompressed luminance channel. Then, by performing a colorization using the matrix C and the RP set, the color image is reconstructed.

B. Formulating the RP Extraction Problem into an L_1 Minimization Problem

The colorization process can be expressed in matrix form as follows:

$$\mathbf{u} = C\mathbf{x}. \tag{6}$$

Here, C represents the matrix which performs a colorization process on $\mathbf{x}$ to obtain the colorized image $\mathbf{u}$. Here, $\mathbf{u}$ is a one dimensional vector of size n, representing the image in raster scan order which has n pixels. The Levin's colorization method can be expressed by (6) with $C = A^{-1}$, where C is a square matrix of size $n \times n$. In the proposed method, C has the size $n \times m$, where m is the size of $\mathbf{x}$, and normally $m < n$. Other colorization methods can also be expressed by (6) using different C matrices.

In the colorization process, C and $\mathbf{x}$ are given and $\mathbf{u}$ is the solution to be sought. In contrast, in colorization based coding, the problem in the encoder is to determine $\mathbf{x}$ when C and $\mathbf{u}$ are given. For the aim of compression, we seek to obtain a sparse vector $\mathbf{x}$. Therefore, we formulate the problem of selecting the RP as an L_0 minimization problem

$$\underset{\mathbf{x}}{\operatorname{argmin}}\, |\mathbf{x}|_0, \text{ s.t. } \mathbf{u}_0 = \mathrm{C}\mathbf{x}. \tag{7}$$

Unfortunately, the matrix C does not satisfy the RIP (Restricted Isotropy Property) condition. Therefore, the problem in (7) cannot be changed directly into an equivalent L_1 minimization problem. To change the L_0 minimization problem in (7) into an equivalent L_1 minimization problem, we have to manipulate the matrix C to satisfy the RIP condition. We multiply a random gaussian matrix R_G, which derives its entries from a zero-mean gaussian distribution, to C (and to $\mathbf{u}_0$) to obtain the matrix $R_G C$ which satisfies the RIP condition. The size of R_G is $l \times n$, where normally $l < n$. This results in the following L_0 minimization problem

$$\underset{\mathbf{x}}{\operatorname{argmin}}\, |\mathbf{x}|_0, \text{ s.t. } \mathrm{R_G}\mathbf{u}_0 = \mathrm{R_G C}\mathbf{x} \tag{8}$$

which can be changed into the following equivalent L_1 minimization

$$\underset{\mathbf{x}}{\operatorname{argmin}}\, |\mathbf{x}|_1, \text{ s.t. } \mathrm{R_G}\mathbf{u}_0 = \mathrm{R_G C}\mathbf{x}. \tag{9}$$

Equation (9) is an optimization problem and can be solved by linear programming such as the Basis Pursuit (BP) or Orthogonal Matching Pursuit (OMP) [6] [7]. Since (9) is an optimization problem, the various results from variational methods can be incorporated into the colorization based coding problem. For example, (9) can also be reformulated as an unconstrained problem

$$\underset{\mathbf{x}}{\operatorname{argmin}}\, |\mathbf{x}|_1 + \lambda \| R_G \mathbf{u}_0 - R_G C\mathbf{x}\|^2 \tag{10}$$

which can be solved by various algorithms developed for unconstrained problems, e.g., the Bregman iterative algorithm [14]. Other weighted L_2 norms that consider the image characteristics can be used instead of the normal L_2 norm in (10).

The unconstrained problem (10) is a more practical choice, since the linear constraint in (9) is too strong a constraint, and therefore, the BP solver will often fall into an unsolvable state. For the unconstrained problem, the solution varies as the parameter λ varies.

In fact, we can alternatively solve the following problem, such that we can control the number of nonzero components as we are minimizing the error between the reconstructed and the original color images

$$\underset{\mathbf{x}}{\operatorname{argmin}}\, \|\mathbf{u}_0 - C\mathbf{x}\|^2, \text{ s.t. } |\mathbf{x}|_0 \le \mathrm{L} \tag{11}$$

where L is a positive integer that controls the number of nonzero components in $\mathbf{x}$. The number L can be determined by the desired compression rate, i.e., if we want a large compression rate, then L is set to a small number. The above problem can be solved either by the BP or the OMP solver.

By formulating the colorization based coding into an L_1 minimization problem, we obtain the following benefits:

1) Compared to the sets of RP obtained by other conventional colorization based coding methods, which are updated at each iteration, the set of RP in our method is obtained only once and requires no update.
2) Compared to other colorization based coding methods, our method needs no extra RP extraction/reduction.
3) It is mathematically guaranteed that the RP set is optimal with respect to the given matrix C in the sense that it minimizes the number of RP due to the L_1 norm. If using (10) or (11), then it is also optimal (with respect to given matrix C) in the sense that it makes the square error in (10) minimum. When solved with the BP/OMP solver, the solution becomes a local optimal minimum of (11).
4) There is no need to adopt geometric methods into the proposed method.
5) By formulating the problem of RP selection as an optimization problem, we have designed a way to adopt existing optimization techniques to the problem of RP selection.

There is one more main difference between the proposed method and other colorization based coding methods. In other colorization based coding methods, the chrominance values of the RP are the same as in the original image $\mathbf{u}_0$. However, in our method the chrominance values of the RP are determined so that they satisfy the constraint in (9) or make the square error in (10) or (11) minimum. Therefore, the chrominance values obtained by the proposed method generally are not the same as those in the original image.

C. Construction of the Colorization Matrix

For the problems in (9)–(11) to be solved, the matrix C has to be determined. In [2]–[4], the matrix C is computed by taking the inverse of A, i.e., $C = A^{-1}$ as has been proposed in [5]. However, in our approach, we cannot let $C = A^{-1}$, since A depends on the set of RP, and therefore, cannot be determined beforehand. In contrast, we want to obtain C directly.

In [4], the observation has been made that the column vectors of the colorization matrix can be regarded as the colorization basis vectors. We construct the matrix C by considering this fact. The chrominance image $\mathbf{u}$ is the linear combination of these column vectors treating the nonzero values (RP) in $\mathbf{x}$ as the coefficients. For $\mathbf{x}$ to become sparse, i.e., to have only a few RP, the column vectors corresponding to the RP should have a large effect on the chrominance image $\mathbf{u}$. Generally, if a column vector in C has many large nonzero entries, this means that the corresponding RP has a large effect on the chrominance image $\mathbf{u}$. In other words, this means that in the decoder, a large region can be colorized by this RP. Therefore, the $\mathbf{x}$ vector obtained by (9)–(11) should contain the RP (or coefficients) which correspond to the column vectors that have large effects on the colorized image. Furthermore, RP corresponding to column vectors with a similar effect on the image should not be redundantly chosen.

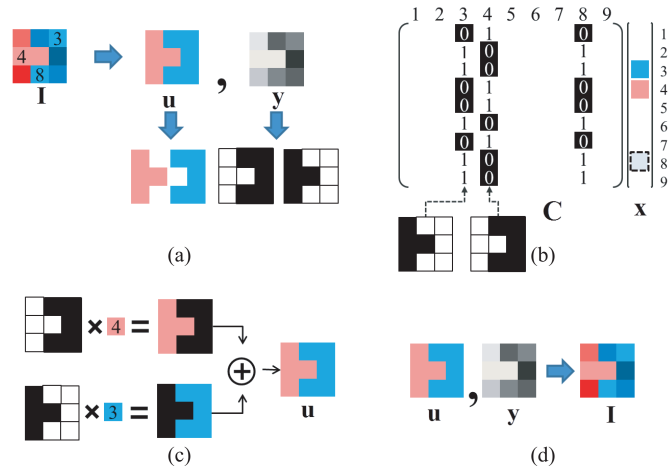


Fig. 2. Exemplary image of reconstructing the colorization matrix. (a) Decomposing the image, (b) constructing the colorization matrix, (c) reconstructing the chrominance channels, and (d) reconstructing the color image.

We explain this using a simple exemplary 3×3 image (image ($\mathbf{I}$) in Fig. 2(a)). We suppose that after decomposing the image ($\mathbf{I}$) into the luminance channel ($\mathbf{y}$) and the color components ($\mathbf{u}$), the color components is mainly constituted of two colors (as can be seen in $\mathbf{u}$ in Fig. 2(a)). This means that the color image can be reconstructed in the decoder using a minimum of two color values and the luminance channel sent from the encoder. Therefore, sending the color information of the third and the fourth pixels could be enough to reconstruct the color image if C sufficiently reflects the effect of the color information on the colorized image. To obtain such a colorization matrix C, segmentation is performed on the luminance channel, which results in two segmented regions corresponding to the two main color components. Then, the matrix C is constructed considering the segmentation result, as in Fig. 2(b). We can observe from the matrix C, for example, that the color information of the third pixel has a full effect on the pixels which positions correspond to those having the value '1' in the third vector, while it has no effect on the pixels which positions correspond to those having zero values. Furthermore, the third and the fourth vectors together have an effect on all the pixels in the image. Therefore, using this C in the RP extraction, the solution vector $\mathbf{x}$ is obtained such that it has only two nonzero values, since a third value would be superfluous. In the decoder, the color components of all the other pixels are recovered using the two nonzero color component values (Fig. 2(c)), and combined with the luminance channel, the color image is reconstructed (Fig. 2(d)). From this simple example, we see that an important step to obtain the matrix C is the segmentation on the luminance channel in the encoder.

D. Segmentation Techniques in the Construction of the Colorization Matrix

As has been seen in the previous section, segmentation plays an important role in the construction of the colorization matrix. In this section, we explain why we have to use a multi-scale segmentation.

1) Meanshift Segmentation: In this paper, we use the meanshift segmentation [12] due to its several desirable properties. The mean shift segmentation uses two parameters where one decides the photometric distances between the pixels inside the segmented regions, and the other decides the spatial distances. Therefore, using the meanshift segmentation, we can easily generate segmented regions of different photometric and spatial characteristics.

Other segmentation techniques may also work with the proposed compression framework if they are tuned to suit well with the proposed method.

2) Multiscale Segmentation: We perform a multi-scale meanshift segmentation to construct the colorization basis. The reason that we use a multi-scale segmentation is that there is the possibility that some regions in the colorized image may lack either the Cb or the Cr components when using a single-scale segmentation. This is due to the fact that even though the RP for both the Cb and Cr components have to be selected for every segmented region, some may not be selected due to the L_1 minimizing constraint.

A multi-scale meanshift segmentation is performed at different scales by using kernels with different bandwidths. A kernel with large bandwidth segments the image into large segments, while a kernel with smaller bandwidth segments the image into smaller segments. This will result in segmented regions of different scales. Now, at each scale, the domain of the whole image (Ω) is the union of the segmented regions

$$\Omega = \Omega_1^k \cup \Omega_2^k \cup \ldots . \Omega_{n_k}^k \ \ \forall k \tag{12}$$

where, k denotes the scale level, n_k is the total number of segmented regions at the k-th scale, and Ω_j^k denotes the j-th segmented region at the k-th scale level. Figure 4 shows the segmentation results and the segmented regions for three different scales of the multi-scale segmentation. Next, we construct a column vector for every segmented region. That is, for every segmented region Ω_j^k, $j = 1, \ldots n^k$ at every scale k, we construct a column vector $\mathbf{c}_{j,k}$ so that the i-th component of the vector $\mathbf{c}_{j,k}$ has a nonzero value if the corresponding i-th pixel lies in the region Ω_j^k, and 0 if not. We then construct the matrix C by putting the column vectors $\mathbf{c}_{j,k,\forall j,\forall k}$ in C, where the column vectors are arranged according to the scale as shown in Fig. 3. Then, the $(i, \sum_{t=1}^{k-1} n_t + j)$-th entry of C becomes

$$C\left(i, \sum_{t=1}^{k-1} n_t + j\right) = \begin{cases} W(d_j^k(i)) & \text{if } i \in \Omega_j^k \\ 0 \quad \text{if } i \notin \Omega_j^k. \end{cases}$$

where n_t represents the total number of segmented regions at scale t, and therefore, $\sum_{t=1}^{k-1} n_t + j$ indicates the j-th column vector at scale k. Here, $W(d_j^k(i))$ is a function of the distance $d_j^k(i)$ between the center of the mass of Ω_j^k and the pixel $i \in \Omega_j^k$. Here, we just let $W(d_j^k(i)) = 1/\sqrt{d_j^k(i)}$. The lower row in Fig. 4 shows the segmented regions of the upper row in Fig. 4 with the weight applied, where a brighter pixel corresponds to a larger weight. Using the weight, applied colorization basis results in colorization with larger weights in the centers of the segmented regions.

Obviously, the minimization process in (10) or (11) will select the RP set corresponding to the column vectors containing large segmented regions with large priority, since the L_2 error grows if they are not chosen. This will colorize the reconstructed image at course scale. Since the RP corresponding to the large-scaled segmented regions are chosen with large priority, the image becomes fully colorized leaving no regions lacking the Cb or Cr components.

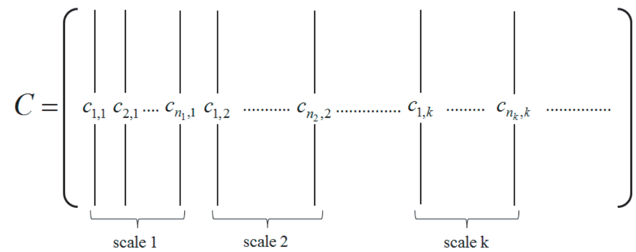


Fig. 3. Colorization matrix C constructed with multiscale segmentation. The colorization basis vectors are arranged according to the scale.

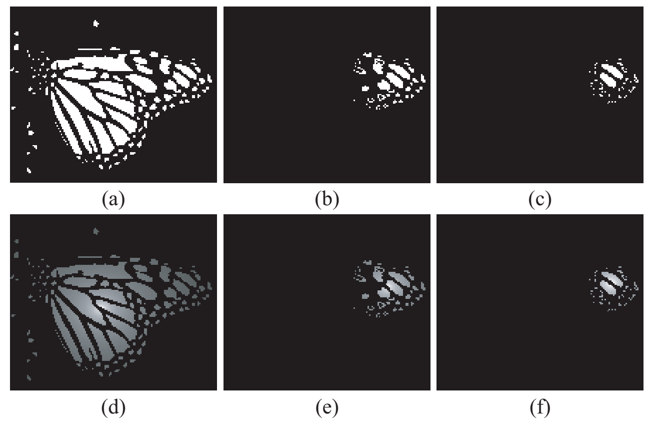


Fig. 4. Segmented regions with different scales obtained by the multi-scale meanshift segmentation. One of the segmented regions of (a) scale 1 $[(h_s, h_r) = (51, 408)]$, (b) scale 5 $[(h_s, h_r) = (25.5, 204)]$, and (c) scale 8 $[(h_s, h_r) = (25.5, 25.5)]$. Showing one of the segmented regions with distance weight of (d) scale 1, (e) scale 5, and (f) scale 8, where a brighter brightness value corresponds to a closer distance to the center of the mass of the segmented region. The brightness values are shown in log-scale for visualty.

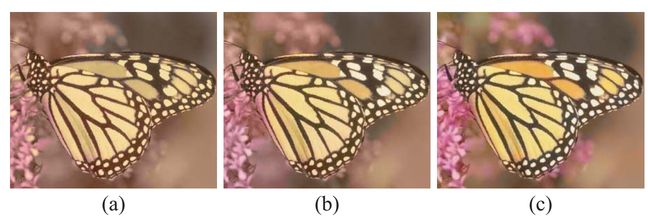


Fig. 5. Colorization results by using the RPs corresponding to (a) scale 1-4, (b) scale 1-8, and (c) scale 1-12, where the scales are as defined in Section 4.

TABLE I
FRAMEWORK OF PROPOSED METHOD

	Description of Proposed Method
Encoder	1) **Input:** Original color image. I 2) **Decomposition:** Decompose I into its luminance channel ($\mathbf{y}$) and original chrominance images $\mathbf{u}^{\mathbf{0}}_{\mathbf{Cb}}$ and $\mathbf{u}^{\mathbf{0}}_{\mathbf{Cr}}$. 3) Perform multi-scale meanshift segmentation on $\mathbf{y}$ to obtain segmented regions Ω^k_j, $j = 1, \ldots n^k$, $\forall_k$. 4) Construct column vectors $\mathbf{c}_{j,k}$ using Ω^k_j, $j = 1, \ldots n^k$, $\forall_k$. 5) Concatenate the column vectors construct $\mathbf{c}_{j,k}$ to construct the colorization matrix C. 6) Using the Orthogonal Matching Pursuit (OMP) obtain the RP set $\mathbf{x}$ by solving the following problem $$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{u}_0 - C\mathbf{x}\|^2, \ s.t. \ \|\mathbf{x}\|_0 \leq L$$ for $\mathbf{u}_0 = \mathbf{u}^{\mathbf{0}}_{\mathbf{Cb}}$ and $\mathbf{u}_0 = \mathbf{u}^{\mathbf{0}}_{\mathbf{Cr}}$. 7) **Output:** Final RP set $\mathbf{x} = \{\mathbf{x}^k_{Cb}, \mathbf{x}^k_{Cr}\}$ (optimal with respect to the matrix C).
Decoder	1) **Input:** Luminance channel ($\mathbf{y}$), RP set $\mathbf{x} = \{\mathbf{x}^k_{Cb}, \mathbf{x}^k_{Cr}\}$. 2) **Reconstruction** 1) Reconstruct the matrix C by performing multi-scale meanshift segmentation on $\mathbf{y}$. 2) Reconstruct the color components by $$\mathbf{u}_{Cb} = C\mathbf{x}_{Cb}$$ and $$\mathbf{u}_{Cr} = C\mathbf{x}_{Cr}.$$ 3) Reconstruct color image I by combining the luminance channel $\mathbf{y}$ and the color components $\mathbf{u}_{Cb}$ and $\mathbf{u}_{Cr}$. 3) **Output:** Reconstructed Color Image I.

After the RP corresponding to large-scaled segmented regions are selected, further RP corresponding to smaller regions are selected according to the errors they reduce. This will add detailed colors to the reconstructed color image in the decoder. Figure 5 shows the reconstructed color images reconstructed with different numbers of RP and different numbers of scales. The different scales are constructed by using different spatial and range parameters in the meanshift segmentation. We explain the details in Section IV. It can be seen that the image reconstructed with a small number of RP is colorized at somewhat course scale, since they correspond to the large-scaled segmented regions, while detailed colors appear in the colorized image when more RP are involved.

Table I shows the framework of the proposed method in pseudo codes. We omitted the compression/decompression process on the luminance channel for simplicity.

IV. Implementation Issues

In this section, we explain the details of the implementation of the proposed method. When constructing the matrix C, we used a 16-scale segmentation, which means that we performed the meanshift segmentation with 16 different spatial and range resolutions, i.e., a combination of four different spatial and four different range resolutions. The parameters h_s and h_r control the spatial and the range resolutions, respectively, and large values of h_s and h_r result in large scaled segmented regions. The meaning of h_s and h_r is the same as in [12]. The parameters (h_s, h_r) used in each scale are as follows: Scale 1: (51, 408), Scale 2: (51, 204), Scale 3: (51, 102), Scale 4: (51, 51), Scale 5: (25.5, 204), Scale 6: (25.5, 102), Scale 7: (25.5, 51), Scale 8: (25.5, 25.5), Scale 9: (15.3, 122.4), Scale 10: (15.3, 61.2), Scale 11: (15.3, 30.6), Scale 12: (15.3, 15.3), Scale 13: (10.2, 81.6), Scale 14: (10.2, 40.8), Scale 15: (10.2, 20.4), Scale 16: (10.2 ,10.2).

The order of the positions of the colorization basis vectors in C is determined by the scale and the mode. That is, the colorization basis vectors are classified first into 16 different parts in C according to the scale. After that, in each part, the colorization basis vectors are ordered again according to the raster scan order of the modes of the corresponding segmented regions. Here, the mode refers to the local maxima of the assumed probability density function of the feature space in the segmented region [12]. Using the same procedure, the same matrix C can be reconstructed in the decoder using the luminance channel sent from the encoder, and therefore, the intended colorized image can be reconstructed using the RP set also sent from the encoder.

We observed that the positions of the RP in $\mathbf{x}$ for the Cb and the Cr components are almost the same, and therefore, we encode the positions only for the Cb components.

V. Experiments

We compared the proposed method with the JPEG and the JPEG2000 standards, as well as two conventional colorization based coding methods, the method of Cheng *et al.* [1] and the method of Ono *et al.* [4]. We used a 4:1:1 color format, which means that the size of the reconstructed Cb and Cr chrominance images are one-fourth of the luminance image. To make the visual comparison easy, we constructed the colors with a very small number of coefficients (or RP) for all the methods. In the comparison with conventional colorization based coding methods, we used an uncompressed luminance channel in the reconstruction of the color image for all methods. The proposed method surpasses other colorization based coding methods by a large amount, and using a compressed luminance channel makes no difference in the comparative result. In the comparison with the JPEG/JPEG2000 standards, we used a compressed luminance channel. Using a compressed luminance channel deteriorates the PSNR a little compared with that using an uncompressed luminance channel.

For conventional colorization based codings, we used 4 bytes to encode each RP, where 2 bytes are used to encode the x and y coordinates, and 2 bytes to encode the Cb and Cr chrominance values. We used a total of 175 RP at the start of the iteration. However, for the method of Ono *et al.*, the number of RP changes after each iteration, and therefore, it was not easy to make the final file size similar to that of ours. For the proposed method, we used (11) with $L = 200$, i.e., we used a total of 200 RP.

We could use more RP than conventional colorization based coding methods, since we need a smaller number of bits to encode the RP. However, the quality of the reconstructed color image is much better with the proposed method even when using the same or even a smaller number of RP. Thus, we use 28 bits to encode each RP, where 12 bits are used to determine the position of the RP in $\mathbf{x}$, and 2 bytes (16 bits)

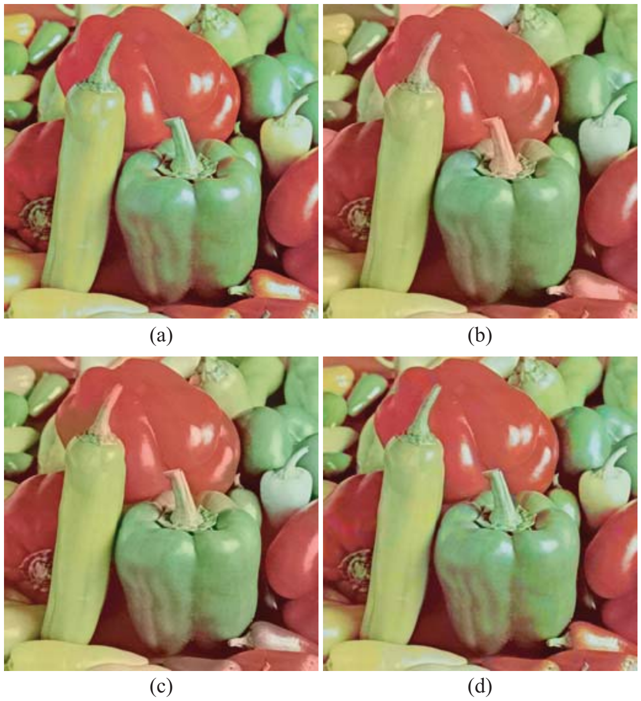


Fig. 6. Experimental results with the 256×256 *"Pepper"* image. (a) Original. (b) Cheng *et al*. (c) Ono *et al*. (d) Proposed.

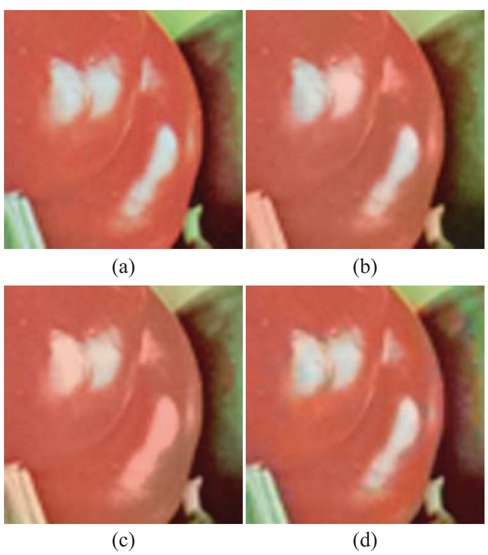


Fig. 7. Showing the enlarged regions of the *"Pepper"* image. (a) Original. (b) Cheng and Vishwanathan [1]. (c) Ono *et al*. [4]. (d) Proposed.

for the chrominance values of the Cb and the Cr components. With 12 bits, we can discriminate between 4096 different basis vectors, which is more than enough, as the number of colorization basis vectors was around 3000 in our experiments.

We experimented with three different color images, two 256 × 256 sized images and one 294 × 250 sized. Figures 6 to 11 compare the reconstructed color images between the different colorization based coding methods.

The PSNR is measured using all the RGB values of the reconstructed color image together, while the SSIM index is measured for the Cb and the Cr channels, separately.

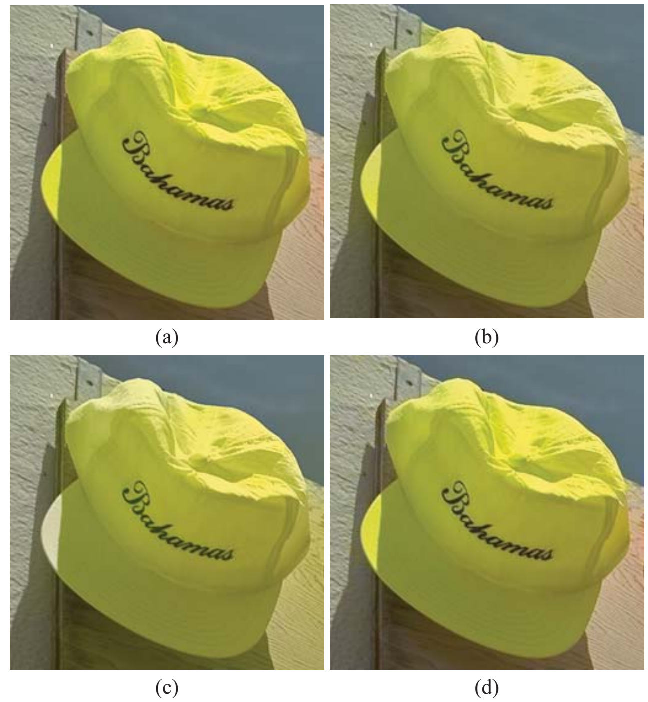


Fig. 8. Experimental results with the 256 × 256 *"Cap"* image. (a) Original. (b) Cheng and Vishwanathan [1]. (c) Ono *et al*. [4]. (d) Proposed.

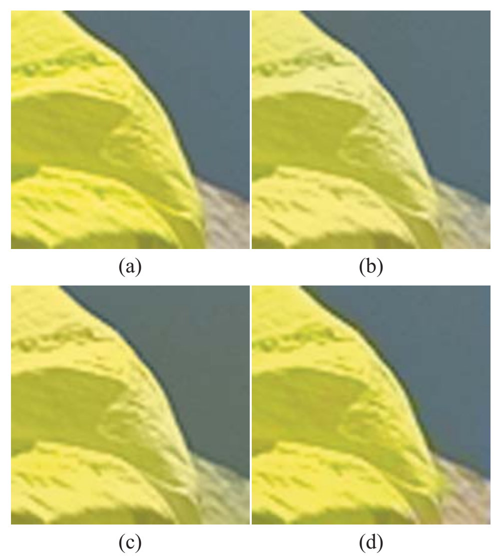


Fig. 9. Showing the enlarged regions of the *Cap* image. (a) Original. (b) Cheng and Vishwanathan [1]. (c) Ono *et al*. [4]. (d) Proposed.

Since all the methods use the same luminance image, we measured the sizes of the files only for the chrominance images. Table II summarizes the sizes of the files, the PSNR, and the SSIM values of the reconstructed images when using different methods. One of the major problems with colorization based coding is the permeation of colors into other color regions. This can be observed in every reconstructed image with other colorization based coding methods. Therefore, the PSNR values are low for other colorization based coding methods. Compared with other colorization based coding methods, the proposed method shows no permeation of colors into other color regions.

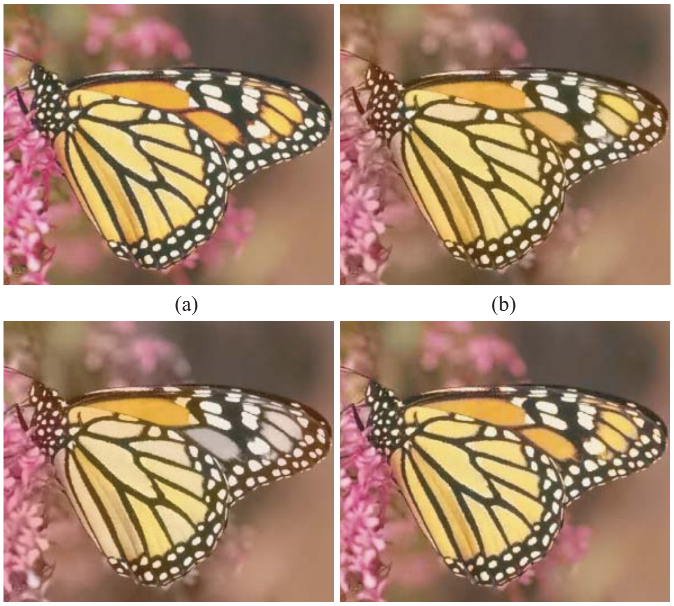


Fig. 10. Experimental results with the 294 × 250 *"Butterfly"* image. (a) Original. (b) Cheng and Vishwanathan [1]. (c) Ono *et al*. [4]. (d) Proposed.

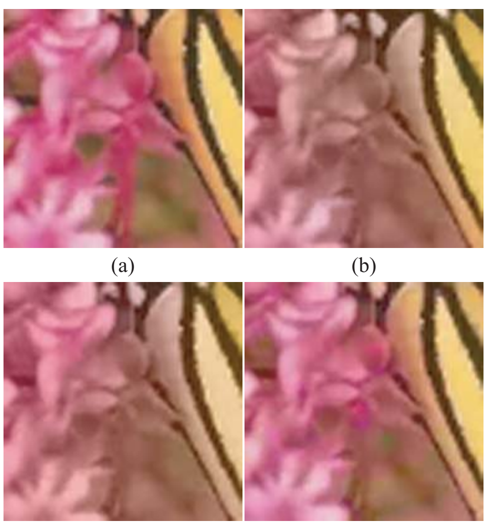


Fig. 11. Showing the enlarged regions of the *Butterfly* image. (a) Original. (b) Cheng and Vishwanathan [1]. (c) Ono *et al*. [4]. (d) Proposed.

For the comparison with the JPEG/JPEG2000 standards, we used standard JPEG/JPEG2000 encoders. We used the JPEG2000 encoder from the Jasper project (www.ece.uvic.ca/∼frodo/jasper/) and the JPEG encoder in the XnView software(www. xnview.com). The file sizes of the images compressed with JPEG/JPEG2000 standards are the sums of the compressed luminance channel and the chrominance values together. With the proposed method, the file size is the sum of the compressed luminance channel and the RP set. Here, we compared the total file sizes of the different methods, and therefore, to match the total file sizes, the sizes of the compressed luminance channels are not the same between the different methods. We further reduced the number of bits used to determine the positions of the RP to 6 bits using delta encoding. This is possible since the distance between each RP does not exceed 64, which fact we verified from experimental results. Therefore, we use a total of 22 bits to encode each RP, where 2 bytes are used to encode the chrominance values. For example, the file size 3.47 KB is obtained by using 0.537 KB (22 bits × 200) for the encoding of the RP and 2.933 KB for the compressed luminance channel.

The reconstructed image using the JPEG standard reveals many blocky artifacts and false colors, which can be observed in the first columns of Figs. 12 and 13, even though the sizes of the files are larger than the reconstructed images using the proposed method. The reconstructed images using the JPEG2000 reveal some ringing artifacts and also some false color artifacts, especially in the regions where the luminance values change a lot. This can be observed especially in Figs. 13(b) and (f). The proposed method surpasses the JPEG standard noticeably in every respect: in the file size, PSNR value, and visual quality. The proposed method is comparable to the JPEG2000 and shows some advantages in some regions. Even though the JPEG2000 standard shows a larger PSNR value for some images, we want to emphasize the fact that the JPEG2000 standard is a highly optimized standard which uses many entropy coding techniques together. If further entropy coding techniques suited for the proposed compression framework are developed, then the performance might surpass the

TABLE II
SUMMARIZATION OF THE COMPARISON OF THE FILE SIZE (KB), PSNR, SSIM VALUES BETWEEN THE DIFFERENT COLORIZATION-BASED CODING METHODS

Image	Method	File Size	PSNR	SSIM (Cb)	SSIM (Cr)
Pepper	Cheng *et al.*	0.7 (KB)	23.49	0.872	0.796
	Ono *et al.*	0.744 (KB)	22.04	0.781	0.757
	Proposed	0.7 (KB)	29.84	0.910	0.973
Cap	Cheng *et al.*	0.7 (KB)	34.37	0.970	0.971
	Ono *et al.*	0.736 (KB)	30.16	0.918	0.923
	Proposed	0.7 (KB)	38.62	0.980	0.994
Butterfly	Cheng *et al.*	0.7 (KB)	26.39	0.789	0.867
	Ono *et al.*	0.904 (KB)	25.17	0.751	0.722
	Proposed	0.7 (KB)	30.93	0.930	0.955

TABLE III
COMPARISON OF THE FILE SIZES (KB) AND PSNR VALUES WITH THE JPEG/JPEG 2000 STANDARDS

Image	Method	File Size	PSNR
Pepper	JPEG	4.22 (KB)	25.7242
	JPEG2000	4.22 (KB)	27.8727
	Proposed	4.08 (KB)	27.3297
	Proposed2	4.08 (KB)	27.6310
Cap	JPEG	3.57 (KB)	28.2731
	JPEG2000	3.59 (KB)	30.0708
	Proposed	3.47 (KB)	30.3384
	Proposed2	3.47 (KB)	30.5175
Butterfly	JPEG	4.80 (KB)	23.8253
	JPEG2000	4.54 (KB)	25.7853
	Proposed	4.54 (KB)	25.6173
	Proposed2	4.54 (KB)	25.9479

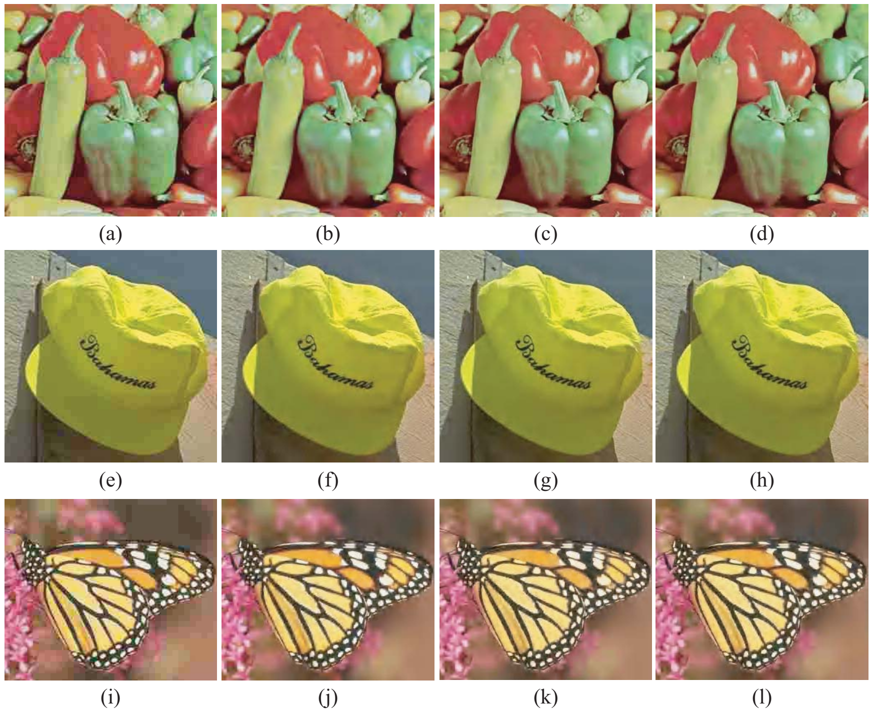


Fig. 12. Reconstructed images using (a), (e), (i) JPEG standard. (b), (f), (j) JPEG2000 standard. (c), (g), (k) Proposed method. (d), (h), (l) Proposed2 method.

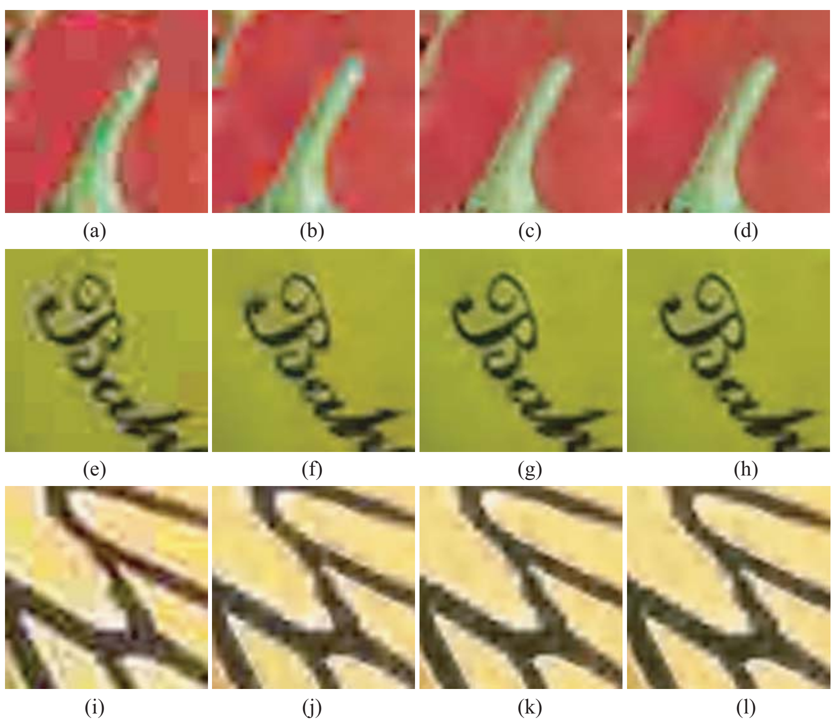


Fig. 13. Enlarged regions of the reconstructed images using (a), (e), (i) JPEG standard. (b), (f), (j) JPEG2000 standard. (c), (g), (k) Proposed method. (d), (h), (l) Proposed2 method.

JPEG2000 standard. To show that the proposed compression framework has potential to further increase its performance, we varied the proposed method a little, by putting some extra small-scaled wavelet basis vectors in the colorization matrix C, together with the basis vectors generated by the meanshift segmentation. This does not increase the file size of the encoded image, due to the fact that wavelet basis vectors can be generated without the knowledge about the image. The chances that the L_2 difference error reduces is now increased, since C contains more column vectors, and (11) will choose the optimal linear combination of the column vectors with respect to the L_2 difference error. Therefore, the PSNR values increases and surpasses the performance of the JPEG2000 as can be seen in Table III, where we denoted the modified proposed method by 'proposed 2'.

The main drawbacks of the proposed method are the high computational cost and use of memory. While the memory could be handled with 2 GB RAM memory, it took about 35 s to construct the C matrix using an Intel CPU with 2.1 GHz speed running under Windows7 as an win32 application.

VI. Conclusion

In this paper, we formulated the colorization based coding problem into an optimization problem. By formulating the problem as an optimization problem we have opened the way to tackle the colorization based coding problem using several well-known optimization techniques. Furthermore, we proposed a method to compute the colorization matrix which can colorize the image with a very small set of RP. Experimental results show that the proposed method surpasses other colorization based coding methods to a large extent in quantitative as well as qualitative measures. The proposed method also surpasses the JPEG standard, and is comparable to the JPEG2000 standard. However, the problem of computational cost and use of large memory remains, and has to be further studied.

References

[1] L. Cheng and S. V. N. Vishwanathan, "Learning to compress images and videos," in *Proc. Int. Conf. Mach. Learn.*, vol. 227. 2007, pp. 161–168.

[2] X. He, M. Ji, and H. Bao, "A unified active and semi-supervised learning framework for image compression," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 65–72.

[3] T. Miyata, Y. Komiyama, Y. Inazumi, and Y. Sakai, "Novel inverse colorization for image compression," in *Proc. Picture Coding Symp.*, 2009, pp. 1–4.

[4] S. Ono, T. Miyata, and Y. Sakai, "Colorization-based coding by focusing on characteristics of colorization bases," in *Proc. Picture Coding Symp.* Dec. 2010, pp. 230–233.

[5] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 689–694, Aug. 2004.

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[7] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[8] E. Candés and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[9] E. Candés, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate information," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1233, 2005.

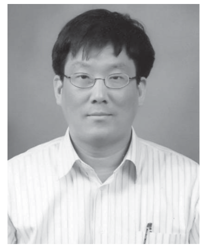
[10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[11] A. Cohen, W. Dahmen, and R. A. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, pp. 211–231, Jun. 2008.

[12] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.

[13] L. Yatziv and G. Sapiro, "Fast image and video colorization using chrominance blending," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1120–1129, May 2006.

[14] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, vol. 1, no. 1, pp. 143–168, 2008.



Sukho Lee received the B.S., M.S., and Ph.D. degrees in electronics engineering from Yonsei University, Seoul, Korea, in 1993, 1998, and 2003, respectively.

He was a Researcher with the Impedance Imaging Research Center from 2003 to 2006 and was an Assistant Professor with Yonsei University from 2006 to 2008. He has been with the Department of Software Engineering, Dongseo University, Busan, Korea, since 2008, where he is currently a Professor. His current research interests include image and video filtering based on PDEs, medical imaging, and computer vision.

Sang-Wook Park received the B.S. and M.S. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2001 and 2003, respectively. He is currently pursuing the Ph.D. degree with Yonsei University.

He was a Researcher with LG Electronics Institute of Technology, Seoul, from 2003 to 2005, and a Senior Researcher with the Electronics and Telecommunications Research Institute, Daejeon, Korea, from 2005 to 2008. His current research interests include image and video denoising, sensor noise modeling and analysis, super-resolution image processing, color interpolation, image restoration, and enhancement.

Paul Oh received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2011.

His current research interests include color interpolation based on local statistics, and image filtering based on sparse approaches.

Moon Gi Kang (M'97) received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1986 and 1988, respectively, and the Ph.D. degree in electrical engineering from Northwestern University, Evanston, IL, USA, in 1994.

He was an Assistant Professor with the University of Minnesota, Duluth, MN, USA, from 1994 to 1997, and since 1997, he has been with the Department of Electronic Engineering, Yonsei University, Seoul, where he is currently a Professor. His current research interests include image and video filtering, restoration, enhancement, and superresolution reconstruction. He has authored more than 100 Technical articles in his areas of expertise.

He served as the Editorial Board Member for the *IEEE Signal Processing Magazine*, the Editor of SPIE Milestone Series Volume (CCD and CMOS imagers), and the Guest Editor of the *IEEE Signal Processing Magazine* Special Issue on Superresolution Image Reconstruction (May 2003). He has served in the technical program and steering committees of several international conferences. He has also served as the Associate Editor of the *EURASIP Journal of Advances in Signal Processing*. He was a recipient of 2006 and 2007 Yonsei Outstanding Research Achievement Awards (Technology Transfer), the 2002 HaeDong foundation Best Paper Award, and the 2000, 2009, and 2010 Awards of Teaching Excellence at Yonsei University.