Color Super-Resolution from a Single CCD

Tomomasa Gotoh   and   Masatoshi Okutomi
Graduate School of Science and Engineering,
Tokyo Institute of Technology
2-12-1, O-okayama, Meguro-ku, Tokyo, Japan
{goto, mxo}@ok.ctrl.titech.ac.jp

Abstract

Limitation in the resolution of CCD image sensors has provided motivation to enhance the resolution of images. Super-resolution has been applied mainly to grayscale images, and producing a high-resolution color image from a single CCD sensor has not been discussed thoroughly. This work aims at producing a high-resolution color image directly from "color mosaic" images obtained by a single-CCD with a color filter array. This method is based on a generalized formulation of super-resolution which performs both resolution enhancement and demosaicking simultaneously. Verification of the proposed method is conducted through experiments using synthetic and real images.

1. Introduction

Each pixel of a single-chip CCD is covered with a color filter. Red, green, and blue are the typical colors used for the filters. The color filters are arranged in a mosaic pattern, and only one primary color is captured for each pixel. The mosaic pattern is called a CFA (Color Filter Array) pattern. Data captured through CFA is thus a color mosaic image (see Figure 7(b)), which is incomplete as a full color image representation.

In order to produce a full RGB image, the missing color channels have to be estimated from the raw data of the color mosaic. This process is generally referred to as "demosaicing".

The simplest demosaicing method is the linear interpolation applied to each color plane. More sophisticated ones [1, 2] have also been reported, which outperform the linear method in color reproduction performance. The major problem in color demosaicing is false colors that occur in the resulting color image. Although low-pass filtering the image reduces the false colors, the resulting image suffers from blurring effect.

The number of pixels contained in each color plane produced by demosaicing is equal to the resolution of CCD. Obtaining a higher resolution is often needed for displaying,

printing, post-processing, etc. The most common solution for such requirement is interpolation.

However, interpolation results in low quality images compared to those captured by higher resolution CCD. This is because no additional information is brought by the interpolation process. Interpolation does not restore the image in detail. In other words, it does not restore any high frequency signal.

This is problematic when displaying an enlarged image on a screen. Display device and capture device often have different standards, and high quality method of resolution conversion is needed in many fields. Applications in computer vision often suffer from bottleneck of resolution, and resolution enhancement is one of the key issues in the challenge of overcoming such limitations.

Super-resolution [3, 4, 5, 6, 7, 8] differs from interpolation in that it restores high frequency details present in the captured scene. Super-resolution is an image processing technique to produce a high-resolution image, from several downsampled images of the same scene with slight motion among them.

Note that super-resolution has been applied to a grayscale image or a full color image, which can be regarded as just a combination of three gray-scale images, and not applicable to raw data obtained from a single-chip CCD. Although sequential processing of demosaicing and super-resolution (Figure 1) results in high-resolution color image, this method suffers from artifacts and blurring effect as seen in the demosaiced images.

The objective of our work is to provide a high-resolution color image reconstruction method using color mosaic images obtained from a single CCD, which is capable of overcoming the limitations described above, utilizing direct access to the raw color mosaic images captured by CFA-masked CCD. The method thus enables effective integration of demosaicing and resolution enhancement.

Unlike a grayscale or 3-CCD camera, a single-CCD camera outputs downsampled color signals, which leads to severe aliasing. Since super-resolution is a method to fuse multiple aliased images to obtain a high-resolution image,
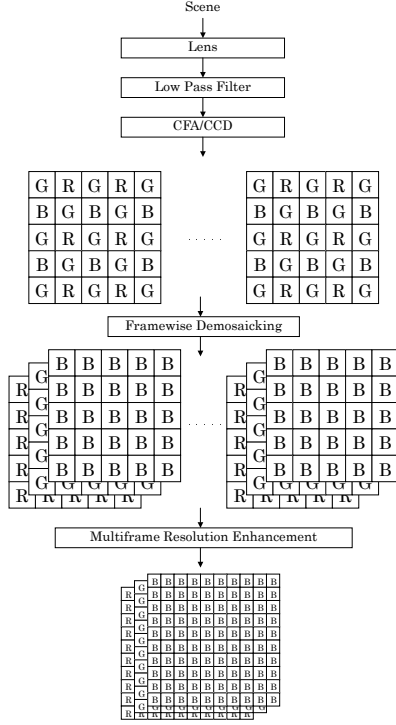
Figure 1. Demosaicing followed by super-resolution.

using raw data is a highly effective way for super-resolution. In addition, eliminating an optical low-pass filter enhances the effectiveness of super-resolution. We thus propose the imaging system configuration illustrated in Figure 2.

In the rest of this paper, Section 2 presents a model for the image reconstruction problem. Section 3 states color image reconstruction as an inverse problem. Experimental results are presented in Section 4, and Section 5 concludes the paper.
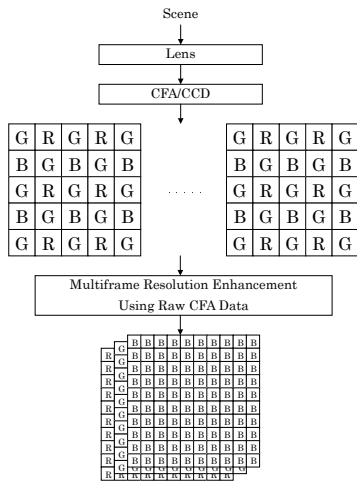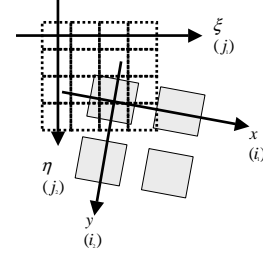

Figure 2. Direct super-resolution from raw data.

Figure 3. Definition of the coordinate system.

2. Observation Model

This section presents the model of single-chip CCD camera generating a sequence of several raw images. We start modeling without considering the effect of CFA. Consider a grayscale image (representing one of the color channels).

The image formation model for a CCD camera can be represented by continuous-discrete model where the input image is continuous while the output data is discrete:

$$u(i_1, i_2) = \iint p(i_1 - x, i_2 - y)I(x, y)dxdy, \quad (1)$$

where $u(i_1, i_2)$ is a digital image obtained by the CCD, and $I(x, y)$ is the true image (scene). (i_1, i_2) is a discrete coordinate system while (x, y) is a continuous coordinate system. $p(x, y)$ is a blurring function known as a point-spread function. PSF introduced due to the optics and the CCD aperture in the system may be approximately modeled by the Gaussian function.

Consider a coordinate transformation $(x, y) = \mathbf{s}(\xi, \eta)$, where (ξ, η) is the coordinate over which the high-resolution image is defined (see See Figure 3). Applying the coordinate transformation, Equation (1) becomes

$$u(i_1, i_2) = \iint p((i_1, i_2) - \mathbf{s}(\xi, \eta))I(x, y) \left| \frac{\partial \mathbf{s}}{\partial(\xi, \eta)} \right| d\xi d\eta. \quad (2)$$

In order to obtain a discrete approximation of the above model, we assume the true image $I(x, y)$ to be constant over the region covering a high-resolution pixel located at high-resolution grid point (j_1, j_2). Then, the integration in Equation (2) can be written in the form

$$u(i_1, i_2) = \sum_{j_1} \sum_{j_2} z(j_1, j_2)h(i_1, i_2, j_1, j_2; \mathbf{s}), \quad (3)$$

where $z(j_1, j_2)$ is the assumed constant value of the true image $I(x, y)$, and

$$h(i_1, i_2, j_1, j_2; \mathbf{s}) = \int_{j_1-1/2}^{j_1+1/2} \int_{j_2-1/2}^{j_2+1/2} p((i_1, i_2) - \mathbf{s}(\xi, \eta)) \left| \frac{\partial \mathbf{s}}{\partial(\xi, \eta)} \right| d\xi d\eta.$$

$$(4)$$

The above integration is performed over the region covering the high-resolution pixel at (j_1, j_2).

Color images can be represented by multiple color channels of 2-D signals. Red, green, and blue are often employed in consumer cameras, such as 3-CCD digital cameras. Let us now extend the imaging model discussed above to color images. Since the model (3) can be considered as a model for one of the color channels, the image formation model for each color channel can be represented as

$$u_c(i_1, i_2) = \sum_{j_1} \sum_{j_2} z_c(j_1, j_2) h(i_1, i_2, j_1, j_2; \mathbf{s}), \quad (5)$$

where $c \in \{R, G, B\}$.

In a single-chip CCD camera, input data is masked by a CFA to produce a color mosaic. Figure 7(b) shows a color mosaic image obtained through a popular CFA known as the Bayer pattern. This process can be modeled as sampling applied to the color image by the CFA.

$$\begin{aligned} y_c(i_1, i_2) &= m_c(i_1, i_2) u_c(i_1, i_2) \\ &= m_c(i_1, i_2) \sum_{j_1} \sum_{j_2} z_c(j_1, j_2) h(i_1, i_2, j_1, j_2; \mathbf{s}), \end{aligned}$$

$$(6)$$

where $m_c(i_1, i_2)$ is a 2-D array containing one or zero for each element. The elements with the value one have sensitivity to the color $c \in \{R, G, B\}$. Suppose the image size is 5×5, the following array show the sampling array for the G channel in the Bayer pattern.

$$m_G(i_1, i_2) : \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad (7)$$

The sampling array for sensors using other CFA pattern can be constructed similarly.

Now, consider a sequence of raw color mosaic images $y_{c,k}(i_1, i_2)$, where $k = 1, ..., M$. Equation (6) can be rewitten using $y_{c,k}(i_1, i_2)$ and corresponding geometric transformation $\mathbf{s}_k$. Representing the equation in matrix-vector form, we obtain

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{z}, \quad (8)$$

where $\mathbf{y}_k$ contains every pixel values of $y_{R,k}(i_1, i_2)$, $y_{G,k}(i_1, i_2)$, and $y_{B,k}(i_1, i_2)$. $\mathbf{z} = \left[\mathbf{z}_R^T, \mathbf{z}_G^T, \mathbf{z}_B^T \right]^T$ contains every pixel values of $z_R(j_1, j_2)$, $z_G(j_1, j_2)$, and $z_B(j_1, j_2)$. $\mathbf{A}_k$ is a matrix specified by $h(i_1, i_2, j_1, j_2; \mathbf{s}_k)$ and $m_c(i_1, i_2)$. Thus Equation (8) is the model relating the high-resolution color image $\mathbf{z}$ and the k-th image $\mathbf{y}_k$ in the low-resolution raw image sequence.
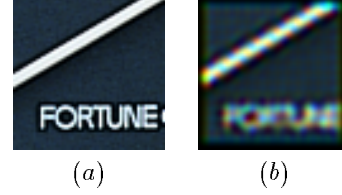

(a) (b)

Figure 4. Artifact in a reconstructed image. High-resolution reference image (a) is used to simulate raw data. Image reconstruction with independent regularization gives (b).

3. Reconstruction of a High-Resolution Color Image from Raw Images

3.1. Problem Formulation

Using forward observation model discussed in the previous section, the image estimate for $\mathbf{z}$ can be obtained by solving the inverse problem.

Using the regularized optimization approach, we have

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \{ f_1(\mathbf{z}) + \lambda f_p(\mathbf{z}) \}, \quad (9)$$

where

$$f_1(\mathbf{z}) = \sum_{k=1}^{M} \| \mathbf{y}_k - \mathbf{A}_k \mathbf{z} \|^2. \quad (10)$$

The parameter λ controls the weight between fidelity to the observed data $f_1(\mathbf{z})$ and the regularization term $f_p(\mathbf{z})$.

The regularization term can be considered as generalization of smoothness constraint that has been employed in grayscale super-resolution.

Using independent regularization term for each channel is inappropriate, because natural images have positive correlation among color planes, and each color plane contains object edges highly correlated and aligned to each other. For example, using independent smoothing term for R,G, and B channel separately leads to color artifact as shown in the Figure 4. The regularization term used here is defined as: $f_p(\mathbf{z}) = \sum_{c \in \{R, G, B\}} \| \mathbf{L} \mathbf{z}_c \|^2$, where $\mathbf{L}$ is the Laplacian operator. A single raw image is used in the experiment. The color artifact occurs due to the similar reason for the artifact seen in the image demosaiced with channel-wise linear interpolation. Independently interpolating each channel with non-coincident subsampling affects correlation among the three channels.

This problem occurs due to unavailability of dense sampling as 3-CCD. If multiple images $\mathbf{y}_1, ..., \mathbf{y}_M, (M > 2)$ with slight motion among them is available, virtually dense sampling is provided , which means less color artifact in a fused image. Obviously, this is almost equivalent to acquiring a 3-CCD image.

Availability of sufficient number of multiple images with different motion among them implies over-determinedness of Equation (10), which leads to reduced color artifact.

Actually, the image estimate still suffers from color artifact even if Equation (10) is over-determined. This is because of channel-independent noise and misregistration of each input image.

In order to suppress color artifact, using color space with less correlated color components is appropriate for defining the regularization term. Consider a color space decomposition into luminance and chrominance components.

$$\begin{aligned} \mathbf{z}_Y &= 0.299\mathbf{z}_R + 0.587\mathbf{z}_G + 0.114\mathbf{z}_B \\ \mathbf{z}_{Cb} &= -0.1687\mathbf{z}_R - 0.3313\mathbf{z}_G + 0.5\mathbf{z}_B \\ \mathbf{z}_{Cr} &= 0.5\mathbf{z}_R - 0.4187\mathbf{z}_G - 0.0813\mathbf{z}_B \end{aligned} \quad (11)$$

Using the above color decomposition, the regularization term is defined by the luminance energy term and the chrominance energy term:

$$f_p(\mathbf{z}) = f_2(\mathbf{z}_Y) + f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr}) \quad (12)$$

Natural images have less variance in chrominance than in luminance, and luminance has the most information of images. In addition, many applications require structures of objects to be presented in high resolution, which means luminance has higher priority to the chrominance. Moreover, the human vision has relatively less sensitivity to change in color.

Based on these considerations, we incorporate edge-preserving smoothing constraint to the luminance energy term, preventing structures of objects from over-smoothing.

The energy term which smoothes along edges and not across them [8] is employed here. i.e. using high-pass operator $\mathbf{P}_d, d \in D$ which evaluates directional smoothness in the direction D :{(horizontal, vertical, two diagonal directions, no direction(isotropic smoothness))} at each pixel, the following energy term is defined

$$f_2(\mathbf{z}_Y) = \sum_{d \in D} ||\mathbf{\Lambda}_d \mathbf{P}_d \mathbf{z}_Y||^2 \quad (13)$$

The matrix $\mathbf{P}_d$ represents convolution operation with a kernel $w(l_1, l_2; \theta)$:

$$\begin{aligned} w(l_1, l_2; \theta) &= G_{xx}(l_1 \cos\theta - l_2 \sin\theta, l_1 \sin\theta + l_2 \cos\theta), \\ G_{xx}(x, y) &= \frac{\partial^2}{\partial x^2} G(x, y). \end{aligned} \quad (14)$$

where $G(\cdot)$ is the 2-D Gaussian function, and θ is the angle corresponding to d. Each element of diagonal matrix $\mathbf{\Lambda}_d$ represents weight on high-pass operation in the direction d at each pixel of $\mathbf{z}_Y$. The weight is determined by detecting orientation and strength of edge at each pixel.

On the other hand, isotropic smoothness is considered for the chrominance energy term. This is because of the

difficulty in estimation of the true edge direction (especially when the number of input images is small and Equation (10) becomes under-determined). Using isotropic high-pass filter $\mathbf{H}$, the following energy term is defined.

$$f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr}) = \lambda_c(||\mathbf{H}\mathbf{z}_{Cb}||^2 + ||\mathbf{H}\mathbf{z}_{Cr}||^2) \quad (15)$$

where λ_c is a weight parameter.

Even if Equation (10) becomes over-determined, error in motion estimation or channel-independent noise brings about color artifact in the image estimate. That is why the chrominance energy term is of great importance for color image reconstruction based on raw data fusion.

In order to suppress color artifact in the image estimate, cut-off frequency of the high-pass filter $\mathbf{H}$ is set equal to the bandwidth of the channel with coarsest sampling in the CFA. Setting larger cut-off frequency is possible when multiple images are available and Equation (10) get close to over-determined, preventing oversmoothing of the chrominance.

The experiment in the next section uses Gaussian high-pass filter whose frequency characteristic is given by:

$$H(u, v) = 1 - \exp(-(u^2 + v^2)/2\sigma_c^2), \quad (16)$$

where (u, v) denote spatial frequency and the standard deviation σ_c is considered as the cut-off frequency.

The chrominance energy term suppresses the misalignment between the color channels. This characteristic can readily be verified with the following experiment. Consider a 1-dimensional analogy for the 2-dimensional edge images shown in Figure 5, whose mathematical model is given by:

$$\begin{aligned} z_R(j) &= erf(j) + 2, \\ z_G(j) &= a_1 z_R(j - Dg) + b_1, \\ z_B(j) &= a_2 z_R(j - Db) + b_2 \end{aligned} \quad (17)$$

where $erf(\cdot)$ is the error function [1] , a_1, a_2, b_1, b_2 are constants. Dg and Db represents the displacement of G and B signals respectively. $Dg = Db = 0$ means perfectly aligned RGB edges which represents two neighboring regions of different colors sharing a common boundary. $Dg \neq 0, Db \neq 0$ means misalignment among RGB signals, witch is a typical cause of color artifact at region boundary.

Changing values of Dg and Db in the model of Equation (17) and calculating the corresponding value of $f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr})$ gives the plot shown in Figure fig6. (The constants are set to $(a_1, a_2, b_1, b_2) = (1.2, 0.8, 0.1, 0.1)$ in this calculation.) Figure 6 shows that the spectral energy function has a minimum when RGB signals are aligned exactly to each other. Such a property of the chrominance energy function holds true for arbitrary a_1, a_2, b_1, b_2, where $a_1, a_2 > 0^2$. This can readily be verified empirically.

---

[1] $erf(\xi) = \frac{2}{\sqrt{\pi}} \int_0^\xi \exp(-t^2)dt$

[2] It holds true, assuming positive correlation among RGB signals.
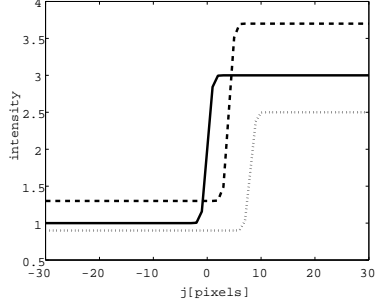
Figure 5. Edge model for red (solid line), green (dashed line), and blue (dotted line) signals.
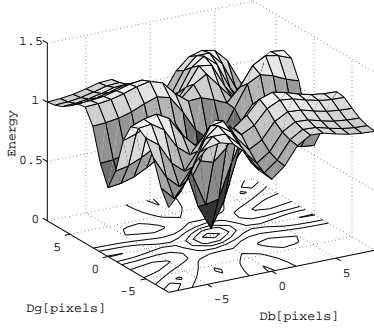

Figure 6. Chrominance energy function characteristic.

3.2. Estimating Motion of Raw Images

The geometric transformation $\mathbf{s}_k$ in the observation model can often be identified prior to super-resolution process.

In the experiments of the following section, we have executed motion estimation with two-parameter model:

$$\mathbf{s}_k(\xi, \eta) = \frac{1}{f} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} d_{xk} \\ d_{yk} \end{bmatrix} \qquad (18)$$

where f is the resolution enhancement ratio, and $\mathbf{d}_k = [d_{xk}, d_{yk}]$ are the parameter to be estimated. First, linear demosaicing (linearly interpolating each color channel independently) is applied to raw input images. Motion estimation is then applied to luminance component of each image. The first image $\mathbf{y}_1$ is the reference image, and the motion of remaining images $\mathbf{y}_2, ..., \mathbf{y}_M$ relative to the reference image was estimated. We have employed the subpixel motion estimation [9], which features practical and precise estimation based on EEC(estimation error cancel).

Super-resolution is a de-aliasing process utilizing subpixel motion in multiple images. It follows that the fractional portion of the motion has the fundamental influence on the image estimate. For example, in super-resolution using grayscale images, using images with motion which has nonzero fractional portion gives much higher quality of image estimate, compared to using images with motion of integer values.

Now, consider a CFA with periodic arrangement of 2x2 kernel such as the Bayer filter. The system matrix $\mathbf{A} = \left[\mathbf{A}_1^T, ...\mathbf{A}_M^T\right]^T$ with a motion $[d_{xk}, d_{yk}]$ and another $\mathbf{A}$ with a motion $[d_{xk} + 2K, d_{yk} + 2L]$($K$,$L$ are arbitrary integers.) have identical structure (neglecting the non-overlapping region of multiple input images). Therefore, the structure of $\mathbf{A}$ has periodicity of 2x2 pixels. Thus, the motion is characterized with the following quantity:

$$\mathbf{r}_k = \mathbf{d}_k \bmod 2 \qquad (19)$$

Input images with their $\mathbf{r}_k$'s uniformly dispersed over the region $[0, 2) \times [0, 2)$ leads to a high quality image estimate, while extremely non-uniform distribution of $\mathbf{r}_k$'s means singularity of $\mathbf{A}$, which leads to low-quality image estimate.

3.3. Implementation

The overall algorithm of the proposed high-resolution color image reconstruction is described here.

(1) Acquire input raw images $\mathbf{y}_1, ..., \mathbf{y}_M$.

(2) Calibrate the CFA model $m_c(i_1, i_2)$, if the CFA arrangement is unknown. (Capturing red, blue, and green objects gives the coresponding pixel sites in the image.)

(3) Produce full color image from input raw data using linear interpolation, and estimate motion $\mathbf{d}_k = [d_{xk}, d_{yk}]$ using luminance component.

(4) Set desired resolution enhancement ratio f.

We now execute optimization of Equation (9). Employing the steepest descent technique gives the following iterative algorithm.

(5) Let $n = 0$. Produce high-resolution color image by interpolating the reference low-resolution image $\mathbf{y}_1$. Let the produced high-resolution image be the initial image estimate $\mathbf{z}^{(0)}$.

(6) Apply edge orientation analysis [8] to the luminance of $\mathbf{z}^{(n)}$, and determine the weight $\mathbf{\Lambda}_d$ in the regularization term $f_2(\mathbf{z}_Y)$.

(7) Update the image $\mathbf{z}^{(n)}$ with the following equations:

$$\mathbf{z}^{(n+1)} = \mathbf{z}^{(n)} - \alpha \left[\frac{\partial}{\partial \mathbf{z}}(f_1(\mathbf{z}) + \lambda f_p(\mathbf{z})) \right]_{\mathbf{z}=\mathbf{z}^{(n)}} \qquad (20)$$

$$\left. \frac{\partial f_1(\mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{z}^{(n)}} = \sum_{k=1}^{M} \mathbf{A}_k^T \left(\mathbf{A}\mathbf{z}^{(n)} - \mathbf{y}_k \right) \qquad (21)$$

$$\left.\frac{\partial f_2(\mathbf{z}_Y)}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}^{(n)}} = \sum_{d \in D} \mathbf{T}_Y^T \mathbf{\Lambda}_d^T \mathbf{P}_d^T \mathbf{P}_d \mathbf{\Lambda}_d \mathbf{T}_Y \mathbf{z}^{(n)} \quad (22)$$

$$\left.\frac{\partial f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr})}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}^{(n)}} = \lambda_c \sum_{C \in \{Cb, Cr\}} \mathbf{T}_C^T \mathbf{H}^T \mathbf{H} \mathbf{T}_C \mathbf{z}^{(n)}$$

$$(23)$$

where α is the step size that controls convergence of the iterative computation. $\mathbf{T}_Y$, $\mathbf{T}_{Cb}$, $\mathbf{T}_{Cr}$ are matrices that operate on $\mathbf{z}$ to produce $\mathbf{z}_Y$, $\mathbf{z}_{Cb}$, $\mathbf{z}_{Cr}$ respectively. (Computation from Equation (20) to Equation (23) do not necessarily require matrix operation, and can be implemented using ordinary image processing including linear filtering, sampling, coordinate transformation, color space transformation, etc.)

(8) Set $n = n + 1$, and go back to (6).

(9) Stop the iteration, if $\mathbf{z}^{(n)}$ has converged or a specified stopping condition has been satisfied.

Note that edge orientation analysis is performed *during* the iteration, and edge direction at each pixel is also updated in the course of iteration. This enables precise detection of edge orientation.

4. Experiments

4.1. Simulated Images

Figure 7(b) shows a simulated raw CFA-masked image, which would have been observed from the scene shown in Figure 7(a). For the experiments here, we assume that the motion parameters are known.

Using the simulated raw images, we have produced a high-resolution image with the proposed direct method. We have also implemented conventional two-pass sequential algorithm (first demosaic and then enhance the resolution). A typical two-pass algorithm would be either of the following algorithms:

(A) single-frame algorithm: First demosaic a raw image and then interpolate to a high-resolution image.

(B) multi-frame algorithm: First demosaic multiple raw images and then apply conventional super-resolution.

The simplest example for the algorithm (A) is linear demosaicing followed by bi-linear interpolation. We first obtain Figure 7(c) and then Figure 7(d). Another example for the algorithm (A) is demosaicing [2] followed by bi-cubic interpolation. We first obtain Figure 7(f) and then Figure 7(g). The resolution is enhanced by factor of two (in both horizontal and vertical directions) in these experiments. Typical color artifacts and blurring effect are contained in these images. These are the major drawbacks of using sequential demosaicing-interpolating process.

We are also interested in the multi-frame algorithm (B), as a comparison with the proposed method. Figure 7(e)

shows a result of linear demosaicing applied to each of multiple raw images, followed by conventional super-resolution [5] applied to each color channel. Similarly, using demosaicing [2] followed by super-resolution [5] gives Figure 7(h).

Figure 7(i)(j)(k) show three restored images by our proposed method, varying in the number of input images M and resolution enhancement ratio f. The resolution enhancement ratio of Figure 7(i) is one and Figure 7(j)(k) is two. The number of input images for Figure 7(i)(j) is one and Figure 7(k) is eight.

Comparing Figure 7(e)(h) to Figure 7(k), we can see that high frequency component is more effectively restored, and color artifacts are restricted in the image estimated with the proposed method.

Quantitative evaluation of the proposed method is also conducted in the simulation for different number of input raw images. For comparison, conventional methods using algorithm (B) is also performed. The resolution enhancement ratio is two in this evaluation. The root mean square(RMS) error between the reference image (Figure 7(a)) and the reconstructed images are evaluated. The RMS error plot is shown in Figure 8. Comparing the proposed method to the existing alternatives, significant improvement in the image reconstruction performance can be seen. The evaluation result also shows contribution of additional input image to the image estimate quality.

4.2. Real Images

We have also conducted the experiment using various real images. Figure 9(1)(2)(3) are results using hand-held camera, and Figure 9(4) show the results using fixed camera capturing objects swinging in the wind. The latter case is not modeled by global motion, and the motion estimation is applied to the local block within the images.

Using as many images as possible do not necessarily guarantee high quality image estimate in experiments using real images. Some reasons for this are that input images may be misregistered, real motion is not properly approximated by 2-parameter model, or illumination may change from image to image. In order to overcome these problems, input images are selectively used with the following priorities: (1)Give high priority to such images that is temporally close to the reference input image. (2)Give high priority to images whose subpixel motion parameters $\mathbf{r}_k$'s are uniformly dispersed over the region $[0, 2) \times [0, 2)$. (3)Remove input images with low matching criteria (SSD, SAD etc).

Figure 9(a) show the first raw image in the captured image sequences. Figure 9(b) show images produced using linear demosaicing followed by bi-linear interpolation. Figure 9(c) show images produced using an existing demosaicing method [2] followed by bi-cubic interpolation. Figure 9(d) shows an image produced by the proposed method.

Figure 7. Reconstructed images. (a)High-resolution reference image. (b)Simulated low-resolution raw image. (c)Linear demosaicing. (d)Linear demosaicing followed by bi-linear interpolation. (e)Linear demosaicing followed by conventional super-resolution [5] (f)Demosaicing [2]. (g)Demosaicing [2] followed by bi-cubic interpolation. (h)Demosaicing [2] followed by conventional super-resolution [5] (i)Proposed, $f = 1, M = 1$. (j)Proposed, $f = 2, M = 1$. (k)Proposed, $f = 2, M = 8$.
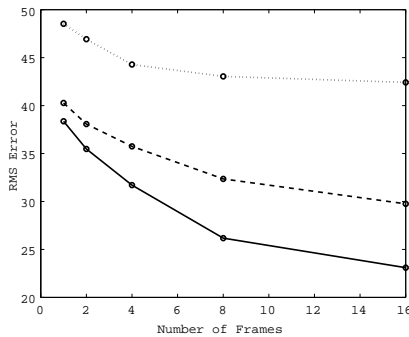

Figure 8. RMS error. dotted line: linear demosaicing followed by conventional super-resolution [5], dashed line: demosaicing [2] followed by conventional super-resolution [5], solid line: proposed.

Resolution enhancement ratio for Figure 9(1)(2)(3) is four, and for Figure 9(4) is two. Effectiveness of the proposed method using real images is also verified in these experiments.

5. Conclusions

The proposed method of high-resolution color image reconstruction has advantage over existing solutions applied to the similar problem. Remarkable effect of the proposed method includes: (a) High quality color imaging: The proposed method provides sophisticated signal processing algorithm, outperforming existing alternatives in the performance of high frequency signal restoration and false color suppression. (b) Direct method: The proposed method is not a mere combination of conventional techniques, consisting of demosaicing and grayscale super-resolution. Sequential processing of the two conventional techniques results in a degraded image. The proposed method is thus a direct method of effective image processing. (c) Post-process: The proposed method does not require changes in capture device itself, and can be implemented as a post-process software. Although there is some conditions of capture device suitable for online super-resolution, images captured in the past can also be used for resolution enhancement.

One of our future work is the motion estimation of raw images. Motion estimation is applied to the demosaiced input images prior to the image reconstruction process in this work. Verification of the preciseness of such a method and the development of more precise and reliable motion estimation method is of great importance. This is because misregistration leads to modeling error of **A**, causing color artifacts in image estimate. Although using severely aliased input images is effective for super-resolution, motion estimation suffers from aliasing. Highly precise and reliable "raw-to-raw registration" is thus an essential factor of super-resolution using raw data.

References

[1] Cok, D.R. Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal. United States Patent 4,642,678, 1987
[2] Laroche, C.A, M.A., Prescott. Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients. United States Patent 5,373,322, 1994
[3] T. S. Huang and R. Y. Tsay. Multiple frame image restoration and registration. in *Advances in Computer Vision and Image Processing*, vol. 1 T. S. Huang, Ed. Greenwich, JAI Press Inc, pp. 317-339, 1984.
[4] M. Irani and S. Peleg. Improving resolution by Image Registration. *CVGIP: Graph. Models Image Process.*, vol. 53, pp. 231-239, Mar. 1991.
[5] R. C. Hardie, K. J. Barnard and E. E. Amstrong. Joint MAP Registration and High-Resolution Image Estimation using a
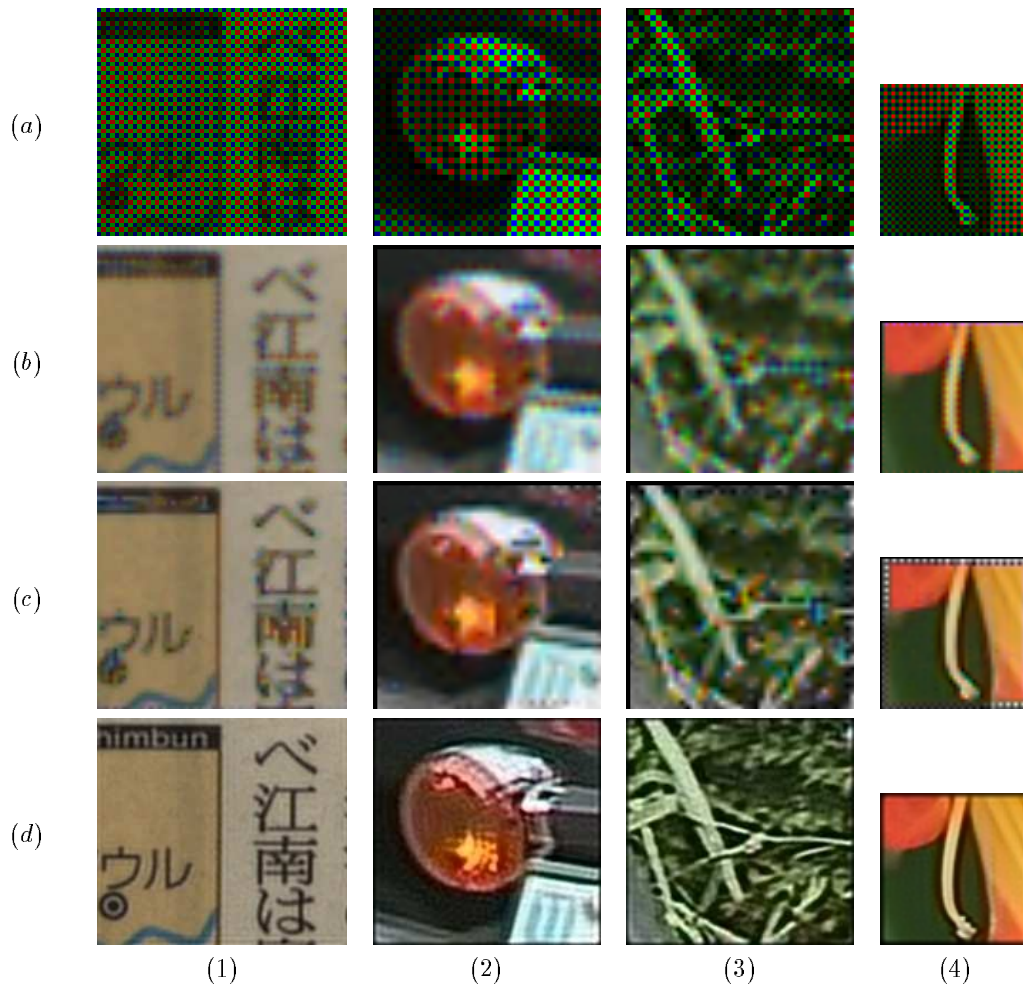
Figure 9. Reconstructed high-resolution images. (a)Observed color mosaic. (b)Linear demosaicing followed by bi-linear interpolation. (c)Demosaicing [2] followed by bi-cubic interpolation. (d)Proposed((1):$f = 4$, $M = 64$, (2):$f = 4$, $M = 64$, (3):$f = 4$, $M = 64$, (4):$f = 2$, $M = 16$.)

Sequence of Undersampled Images. *IEEE Trans. on Image Processing*, vol. 6, pp. 1621-1633, 1997

[6] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan. High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration. *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. III, pp. 169-172, Mar. 1992.

[7] R. R. Schultz and R. L. Stevenson. Improved definition video frame enhancement. *IEEE Int. Conf. Acoustics, Speech, and Signal Processing(ICASSP)*, vol. IV, pp. 2169-2172, May 1995.

[8] J. Shin, J. Paik, J. R. Price, and M.A. Abidi. Adaptive regularized image interpolation using data fusion and steerable constraints. In *SPIE Visual Communications and Image Processing*, volume 4310, pp. 798-808, January 2001.

[9] M. Shimizu and M. Okutomi. Precise sub-pixel estimation on area-based matching. In *Proc. 8th International Conference on Computer Vision*, vol. I, pp. 90-97, July 2001.