

# 単板カラー撮像素子の RAW データを利用した高精細画像復元

後藤 知 将<sup>†</sup>, 奥 富 正 敏<sup>†</sup>

撮像素子を用いた画像センサの物理的分解能の限界を超える高精細画像を生成する方法が研究されてきた。従来の、複数画像から高解像度の画像を合成する超解像法は、主にモノクロ画像を対象としていたため、近年の急速な普及をみる単板カラー撮像素子を用いて高解像度画像を生成する効果的な方法が望まれている。そこで、本論文では単板撮像素子の RAW データを直接利用し、高解像度カラー画像を生成する手法を提案する。本手法は単板カメラの demosaicking 処理を内包する一般化されたカラー超解像問題の定式化を特徴とする。実験では合成画像と実画像を用いて、提案手法に基づく信号処理の有効性を検証する。

## High Resolution Color Image Reconstruction Using Raw Data of a Single Imaging Chip

TOMOMASA GOTOH<sup>†</sup> and MASATOSHI OKUTOMI<sup>†</sup>

The limited resolution of image sensors has motivated the enhancement of image resolution. Super-resolution has been applied mainly to grayscale images, but producing a high-resolution color image using a single-chip imaging device has not been investigated thoroughly. This work aims at producing a high-resolution color image directly from raw data obtained by a single imaging chip employing a color filter array. This method is based on a generalized formulation of super-resolution that simultaneously performs both resolution enhancement and demosaicing. The proposed method is verified through experiments using synthetic and real images.

### 1. はじめに

単板カラーカメラにおいて、撮像素子の各要素にはカラーフィルタがかけられている。そのためカラーフィルタの配列 (CFA: Color Filter Array) を通して得られるデータはカラーモザイク状の画像であり、フルカラー画像の表現としては不完全なものとなる。

完全なカラー画像を生成するには、欠けている色チャンネルをカラーモザイク状の RAW データから推定する必要がある。一般に、この色復元の処理を demosaicking と呼ぶ。最も簡単な demosaicking は、各色チャンネルごとに線形補間を行うことである。より高度な demosaicking 手法<sup>1),2)</sup> も報告されており、線形補間と比較して画像復元の精度が高い。

一般に、demosaicking において、処理結果の画像に偽色が発生することが主な問題となる。光学ローパスフィルタに通すことにより偽色を低減させることは

できるが、その結果帯域の制限を受けることになる。特に、画像を高解像度の画面に拡大表示する場合、帯域制限による画質の低下が顕著となる。通常、画像の撮像素子と表示系の解像度やフォーマットは異なる場合が多いことから、このような画質劣化が問題となる場面が多く、そのため高品質な解像度変換手法に対するニーズは非常に高い。コンピュータビジョンのアプリケーションでも解像度の不足が性能のボトルネックとなることがあり、効果的な高解像度化手法が望まれている。

そこで、本研究では複数の入力画像を合成して高解像度画像を生成する超解像<sup>3)~7)</sup>の理論に基づき、単板撮像素子の RAW データ出力から高精細カラー画像を生成する手法を提案する。これまでに超解像はグレースケール画像や三板式撮像素子の出力画像等のカラー画像に適用されてきた。一方、単板撮像素子の RAW データ出力に対しては、まず demosaicking 処理を施した後、得られたカラー画像に超解像を適用する方法も考えられる (図 1)。

しかし、本研究ではこのような 2 段階の処理とは異なる直接的な手法 (図 2) を提案し、その有効性を検

<sup>†</sup> 東京工業大学大学院  
Graduate School, Tokyo Institute of Technology  
現在、ソニー株式会社  
Presently with Sony Corporation

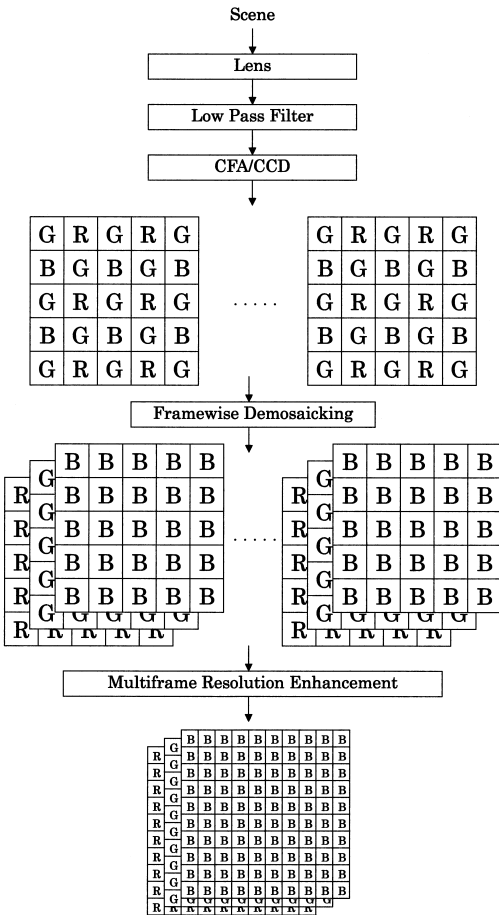


図1 demosaicking および超解像  
Fig.1 Demosaicking and super-resolution.

証する．単板式カラーカメラはモノクロカメラや三板式カメラとは異なり，各色チャンネルがダウンサンプルされた疎なデータを出力するため，エイリアシングが強く入ったデータが得られる．超解像は複数のエイリアシングが入った画像を統合することで解像度を高める手法であることから，RAW データを直接利用することによる効果は非常に高くなる．また，光学ローパスフィルタも極力使用しない図2のような構成が望ましいと考えられる．

本論文の構成は次のとおりである．2章では，単板式カラーカメラによる複数画像の撮像モデルを導出し，そのモデルに基づき3章では高解像度カラー画像の復元問題を定式化する．4章では合成画像と実画像を用いた実験を行い，提案手法の有効性を検証する．5章では，本研究を総括する．

2. 観測モデル

デジタルカメラによるグレースケール画像の撮像

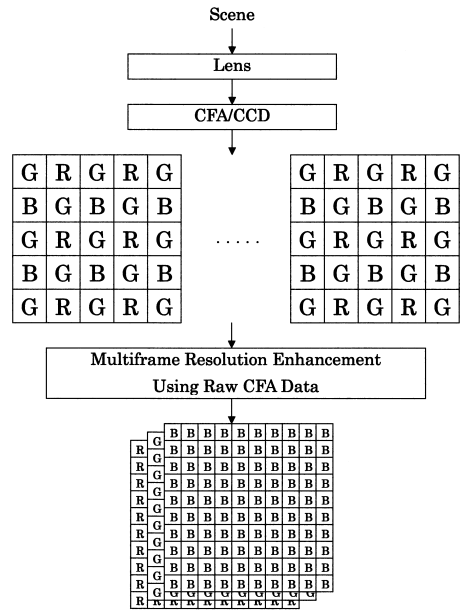


図2 RAW データからの超精細カラー画像の復元  
Fig.2 High resolution color image reconstruction from RAW data.

モデルは，次の連続-離散モデル

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

で与えられる． $u(i_1, i_2)$  は撮像素子から出力されたデジタル画像を表し， $I(x, y)$  は撮像装置の入力となる連続画像を表す． $p(x, y)$  は点広がり関数である．点広がり関数は，光学系と CCD 開口からなる光学的伝達特性から導出されるが，実用上は2次元ガウス関数で近似できる．図3に示す入力画像座標系  $(x, y)$  と復元の対象となる高解像度画像座標系  $(\xi, \eta)$  の間の変換

$$(x, y) = s(\xi, \eta) \quad (2)$$

を考えると，式(1)は

$$u(i_1, i_2) = \iint p((i_1, i_2) - s(\xi, \eta)) I(x, y) \left| \frac{\partial s}{\partial(\xi, \eta)} \right| d\xi d\eta \quad (3)$$

となる．変換  $s$  は射影変換等の一般的な幾何学的変換を表す．

ここで，座標系  $(\xi, \eta)$  の格子点  $(j_1, j_2)$  に位置する高解像度画素を覆う領域  $[j_1 - 1/2, j_1 + 1/2] \times [j_2 - 1/2, j_2 + 1/2]$  全体で  $I(x, y)$  が一定値  $z(j_1, j_2)$  をとると仮定すると，式(3)は次の離散-離散モデル

$$u(i_1, i_2) = \sum_{j_1} \sum_{j_2} z(j_1, j_2) h(i_1, i_2, j_1, j_2; s) \quad (4)$$

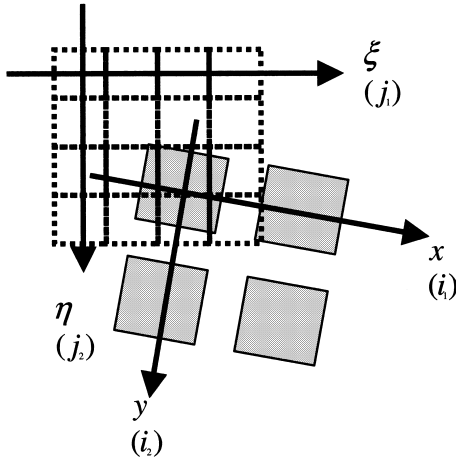


図 3 座標系の定義

Fig.3 Definition of the coordinate system.

となる。ここで、

$$h(i_1, i_2, j_1, j_2; \mathbf{s}) = \int_{j_1-1/2}^{j_1+1/2} \int_{j_2-1/2}^{j_2+1/2} p((i_1, i_2) - \mathbf{s}(\xi, \eta)) \left| \frac{\partial \mathbf{s}}{\partial (\xi, \eta)} \right| d\xi d\eta \quad (5)$$

である。

次に、カラー画像の色チャンネル  $c \in \{R, G, B\}$  の撮像モデルを考えると、チャンネル  $c$  の撮像モデルは式 (4) と同様にモデル化でき、

$$u_c(i_1, i_2) = \sum_{j_1} \sum_{j_2} z_c(j_1, j_2) h(i_1, i_2, j_1, j_2; \mathbf{s}) \quad (6)$$

となる。ここで、色収差の影響は無視し、各色チャンネルの点広がり関数が同一であるという条件を仮定したが、これは実用上妥当な仮定であると考えられる。

上記撮像モデルは三板式カメラモデルに相当するが、単板カメラではさらに CFA の影響も考慮する必要がある。撮像素子にカラーフィルタがかけられた結果、ある画素  $(i_1, i_2)$  において、1 つのチャンネル  $c$  のみに対し感度を持つカラーモザイク画像  $y_c(i_1, i_2)$  が生成される。

$$\begin{aligned} y_c(i_1, i_2) &= m_c(i_1, i_2) u_c(i_1, i_2) \\ &= m_c(i_1, i_2) \sum_{j_1} \sum_{j_2} z_c(j_1, j_2) h(i_1, i_2, j_1, j_2; \mathbf{s}) \end{aligned} \quad (7)$$

ここで、 $m_c(i_1, i_2)$  はカラーフィルタの配列パターンを表す感度配列であり、画素  $(i_1, i_2)$  においてチャンネル  $c$  に感度を持つとき、 $m_c(i_1, i_2) = 1$  であり、その他の画素では  $m_c(i_1, i_2) = 0$  である。たとえば、画像

サイズが  $5 \times 5$  の Bayer パターン<sup>10)</sup> を表す感度配列は次のように書ける。

$$m_R(i_1, i_2) : \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \quad (8)$$

$$m_G(i_1, i_2) : \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad (9)$$

$$m_B(i_1, i_2) : \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (10)$$

次に、複数画像の生成モデルを考える。式 (7) について、 $M$  枚の低解像度画像  $y_{c,k}(i_1, i_2)$ ,  $k = 1, \dots, M$  がそれぞれ互いに異なる変換  $s_k$ ,  $k = 1, \dots, M$  をともなって生成されるモデルに書き換えると、次の画像列生成モデルが得られる。

$$y_{c,k}(i_1, i_2) = m_c(i_1, i_2) \sum_{j_1} \sum_{j_2} z_c(j_1, j_2) h(i_1, i_2, j_1, j_2; \mathbf{s}_k) \quad (11)$$

上記モデルは以下のマトリクス表記で表現できる。

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{z} \quad (12)$$

ここで  $\mathbf{y}_k$  は  $y_{R,k}(i_1, i_2)$ ,  $y_{G,k}(i_1, i_2)$ ,  $y_{B,k}(i_1, i_2)$  のうち、 $m_c(i_1, i_2) = 1$  となる全画素値を含むベクトルである。 $\mathbf{z} = [z_R^T, z_G^T, z_B^T]^T$  は  $z_R(j_1, j_2)$ ,  $z_G(j_1, j_2)$ ,  $z_B(j_1, j_2)$  の全画素値を含むベクトルである。 $\mathbf{A}_k$  は  $h(i_1, i_2, j_1, j_2; \mathbf{s}_k)$  と  $m_c(i_1, i_2)$  で定まる行列である。したがって、式 (12) は、高解像度カラー画像  $\mathbf{z}$  と低解像度 RAW 画像列中の  $k$  番目の画像  $\mathbf{y}_k$  の関係を表す撮像モデルとなる。

### 3. 高解像度カラー画像の合成

#### 3.1 問題の定式化

前章の順方向観測モデル (12) を元に、逆問題解析を行うことにより高解像度画像を推定する。正則化の項を付加した目的関数の最適化手法を用いると、推定すべき高解像度カラー画像  $\hat{\mathbf{z}}$  は次式により与えられる。

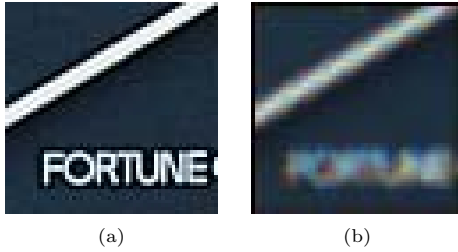


図 4 復元画像の偽色の発生．高解像度の参照画像 (a) からシミュレーションにより RAW データを生成し，独立な正則化項を用いて画像を復元した結果 (b)

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

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \{f_1(\mathbf{z}) + f_p(\mathbf{z})\} \quad (13)$$

ここで，目的関数の第 1 項

$$f_1(\mathbf{z}) = \sum_{k=1}^M \|\mathbf{y}_k - \mathbf{A}_k \mathbf{z}\|^2 \quad (14)$$

は測定データへの適合度を表す項である．目的関数の第 2 項は正則化項である．正則化項は，グレースケール画像の超解像で用いられてきた，信号の局所的な平滑性を評価する項をカラー信号に拡張した評価式と考えることができる．

一般に，自然画像では RGB 信号の各成分は強い正の相関を持ち，エッジ位置も一致していることが知られている．そのため，各チャンネルに対し独立な評価式を  $f_p(\mathbf{z})$  として用いることは望ましくない．たとえば RGB 信号のそれぞれに対して (独立に) 平滑化の効果を持つ評価式を用いると，図 4 (b) のような色ずれが生じる．

この現象は，線形 demosaicking を施した画像に見られる偽色と同様な原因で生じる．すなわち，異なるサンプリング位置を持つ各チャンネルを独立に補間することで RGB 信号間の相関関係を大きく変化させることが原因となる．このことは，RAW データの各チャンネルが三板式撮像素子のように密なサンプリングではないことに起因している．したがって，式 (14) において複数の位置がずれた入力画像  $\mathbf{y}_1, \dots, \mathbf{y}_M$ , ( $M > 2$ ) が得られれば，密なサンプリングをすることと同等の効果が得られ，色ずれの発生は低減される．これは，事実上三板式撮像素子の出力画像を得ていることとほぼ同等であることから明らかである．複数の位置がずれた入力画像が十分に得られ，式 (14) が over-determined に近づくと色ずれが低減されることに

なる．

しかし実際には，式 (14) が over-determined であっても偽色が発生することが確認されている．その原因となるのは，各チャンネルに独立のノイズが入ることや，入力画像間のモーション  $\mathbf{s}_k$  の推定が不正確であることである (なお，本研究では色収差の影響による色ずれは考慮していない)．

このような偽色の発生を抑えるには，RGB より相関の小さい色成分を持つような色空間において  $f_p(\mathbf{z})$  を定義することが適切である．そこで，次のような RGB 空間から YCbCr 空間への変換を考える．

$$\begin{aligned} \mathbf{z}_Y &= 0.299\mathbf{z}_R + 0.587\mathbf{z}_G + 0.114\mathbf{z}_B \\ \mathbf{z}_{Cb} &= -0.1687\mathbf{z}_R - 0.3313\mathbf{z}_G + 0.5\mathbf{z}_B \\ \mathbf{z}_{Cr} &= 0.5\mathbf{z}_R - 0.4187\mathbf{z}_G - 0.0813\mathbf{z}_B \end{aligned} \quad (15)$$

上記変換を用い，正則化項は輝度成分と色差成分に関する項で定義される．

$$f_p(\mathbf{z}) = f_2(\mathbf{z}_Y) + f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr}) \quad (16)$$

自然画像では，輝度成分と比較して色差成分の変化が小さく，輝度成分に最も多くの情報が含まれる．また，多くのアプリケーションでは，色彩の情報以上に輝度情報で表される被写体の構造の方が高精細に表現されることが要求される．さらに，人間の視覚系が色彩の変化に対して比較的感度が低いといった理由から，次のような条件を正則化の項に取り込む．

まず，輝度に対してはエッジ保存型の非等方的平滑性条件を正則化の項に組み込むことで，被写体の構造が過度な平滑化により劣化することを防ぐ．このような要求を満たす平滑性条件として，エッジに沿う方向に平滑化する正則化演算子<sup>8)</sup> を利用する．すなわち，各画素における方向  $D: \{ \text{水平, 垂直, 対角方向 (2通り)}, \text{方向なし (全方向均等)} \}$  への平滑性を評価するハイパス演算子  $\mathbf{P}_d, d \in D$  を用いて次式で表される．

$$f_2(\mathbf{z}_Y) = \sum_{d \in D} \|\Lambda_d \mathbf{P}_d \mathbf{z}_Y\|^2 \quad (17)$$

対角行列  $\Lambda_d$  の各成分は  $\mathbf{z}_Y$  の各画素におけるエッジ方向  $d$  のハイパス演算に対する重みを表す．この重みは，各画素におけるエッジの方向と強度を検出することによって定めることができる．

一方，色差に対する正則化項は，等方的平滑性を考慮する．入力画像の枚数が少なく，式 (14) が未知数に対して under-determined になる場合，色差に対して異方的平滑性条件を与えると，真のエッジ方向の推定が困難となるためである．等方的ハイパスフィルタ  $\mathbf{H}$  を用いた次の正則化項を定義する．

$$f_3(\mathbf{z}_{Cb}, \mathbf{z}_{Cr}) = \lambda_c (\|\mathbf{H} \mathbf{z}_{Cb}\|^2 + \|\mathbf{H} \mathbf{z}_{Cr}\|^2) \quad (18)$$

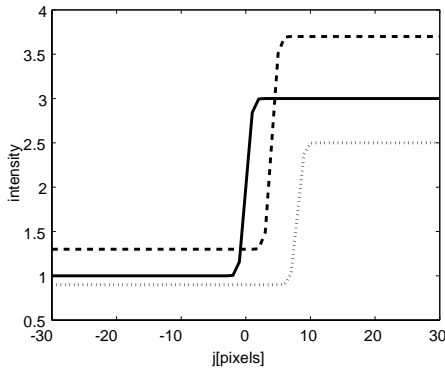


図 5 エッジモデルの R 成分 (実線), G 成分 (破線), B 成分 (点線)

Fig. 5 Edge model for red (solid line), green (dashed line), and blue (dotted line) signals.

$\lambda_c$  は重みパラメータである.

式 (14) が over-determined の場合でも, モーション推定の誤差やチャンネル独立のノイズに起因する偽色が発生するため, 上記の正則化項はカラー画像の復元において重要な役割を果たす.

ハイパスフィルタ  $H$  のカットオフ周波数は, カラーフィルタのサンプリング間隔が最も粗いチャンネルの帯域に設定することで偽色を抑制することができる. 複数の入力画像が与えられ, 式 (14) が over-determined に近づく場合にはより大きなカットオフ周波数を設定することで, 色差の過剰な平滑化を抑制できる. なお, 次章の実験では, 正則化演算子  $H$  としてガウシアンハイパスフィルタを用い, その標準偏差をカットオフ周波数としている.

上述の色差に対する正則化項が RGB チャンネル間のエッジのずれを抑制する性質を持つことは, 次の実験により確かめられる. まず, 画像中で 2 つの色の領域が接しているエッジ部分のモデルとして, 次のような 1 次元の RGB 信号を考える.

$$\begin{aligned} z_R(j) &= erf(j) + 2, \\ z_G(j) &= a_1 z_R(j - Dg) + b_1, \\ z_B(j) &= a_2 z_R(j - Db) + b_2 \end{aligned} \quad (19)$$

$erf(\cdot)$  は誤差関数 を表す.  $a_1, a_2, b_1, b_2$  は定数である.  $Dg, Db$  はそれぞれ R 成分に対する G 成分, B 成分のエッジ位置のずれ量を表す (図 5). RGB 間のエッジ位置が揃っている場合 ( $Dg = Db = 0$ ) は 2 つの色の領域が接しているモデルとなるが, エッジ位

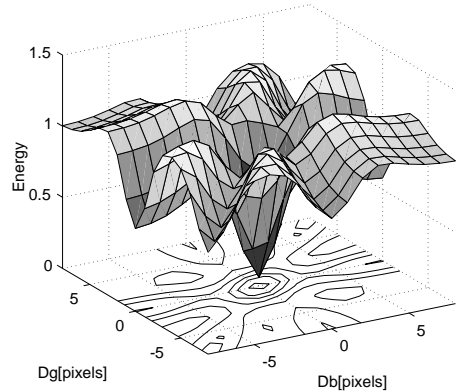


図 6 色差に関するエネルギー関数の性質

Fig. 6 Chrominance energy function characteristic.

置が微小にずれている場合 ( $Dg \neq 0, Db \neq 0$ ) は境界部に色ずれを生じることになる.

モデル (19) の  $Dg, Db$  を変化させたときの  $f_3(z_{Cb}, z_{Cr})$  の評価値の変化を図 6 に示す (定数は,  $(a_1, a_2, b_1, b_2) = (1.2, 0.8, 0.1, 0.1)$  とした). 図 6 から, エッジの位置が完全に一致したときに正則化項の評価値が最小値をとることが分かる. このような正則化項の性質は,  $a_1, a_2 > 0$  なる任意の  $a_1, a_2, b_1, b_2$  に対して成り立つことが容易に確認できる.

### 3.2 モーション推定

目的関数 (13) において, 複数の入力画像間のモーションを表す変換  $s_k, k = 1, \dots, M$  を推定する必要がある. 過去の超解像の文献には並進運動や回転運動等の単純なモデルを用いているものや, 射影変換等の自由度の高いモデルを用いているもの<sup>9)</sup>がある. モーションの自由度は単に高めればよいというものではなく, 場合によってはかえって推定を困難にし, 誤差が大きくなることもあるため, 状況に応じて必要十分な自由度のモデルを採用する必要がある. 本研究では, モーション推定法の検討が趣旨ではないため, 単純なモーションモデルとして, 水平・垂直方向の並進運動を考慮した 2 パラメータの推定を行った. 今回実験した, 手ぶれ程度の微小なモーションの場合には, 並進運動だけでも, 十分な超解像の効果が得られている.

低解像度画像  $y_k$  から高解像度画像  $z$  への倍率を  $f$  とすると, 平行移動は変換

$$s_k(\xi, \eta) = \frac{1}{f} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} d_{xk} \\ d_{yk} \end{bmatrix} \quad (20)$$

で表される. この変換において, パラメータ  $d_k =$

周波数特性が  $H(u, v) = 1 - \exp(-(u^2 + v^2)/2\sigma_c^2)$  で与えられるフィルタ.

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

RGB 信号が正の相関を持つとき成り立つ.

$[d_{xk}, d_{yk}]^T$  を推定した。推定方法は、まず入力 RAW データ  $y_1, \dots, y_M$  に対し線形 demosaicking (R, G, B の各成分を独立に線形補間) を施し、次いで輝度成分を抽出する。この輝度成分に対し、高精度サブピクセルマッチング手法<sup>(11)~(13)</sup> を適用することにより、入力画像のモーションを求めた。本研究では、1 枚目の入力画像  $y_1$  を基準画像とし、 $d_1 = [0, 0]^T$  としている。

超解像法は、各入力画像のサブピクセルの位置ずれを利用する de-aliasing 処理であることから、基準画像に対する位置ずれの小数部分の値が画像の復元精度に本質的な影響を与える。たとえばグレースケール画像では、位置ずれ量が整数となる入力画像列を用いるより、位置ずれ量が小数部分を持つ場合の方が、画像復元精度が高い。

Bayer フィルタのような 2 画素周期のカラーフィルタを用いた場合は、入力画像列の中から、 $[d_{x1}, d_{y1}]^T, \dots, [d_{xM}, d_{yM}]^T$  なる位置ずれを持った画像を選択したときの  $A = [A_1^T, \dots, A_M^T]^T$  と、 $[d_{x1} + 2K_1, d_{y1} + 2L_1]^T, \dots, [d_{xM} + 2K_M, d_{yM} + 2L_M]^T$  ( $K_k, L_k (k = 1, \dots, M)$  は任意の整数) なる位置ずれを持つ画像を選択したときの  $A$  は等しくなる (画像の終端で複数の入力画像が重なりを持たない領域については考慮しない)。したがって、 $A$  の構造は位置ずれに対し 2 画素の周期性を持つため、次の量

$$\text{mod}(d_1, 2), \text{mod}(d_2, 2), \dots, \text{mod}(d_M, 2) \quad (21)$$

が決定すると復元画像  $z$  が一意に決定される。これらの値が領域  $[0, 2) \times [0, 2)$  内で均一に分布するような入力画像列が与えられた場合に良好な復元画像となる。逆に、極端に不均一な分布になる場合は行列  $A$  が特異となり、復元結果の解像感が低くなる。したがって、入力画像列の中から画像を数枚選択して超解像処理を適用するときは、位置ずれの均一性を考慮した選択基準を設けることが望ましい。

なお、本研究では、入力画像から輝度成分を算出してあらかじめ画像のマッチングを行っているが、そのような手法の精度の検証とともに、より高精度で信頼性の高いマッチング手法の開発も必要となる可能性がある。各入力画像間のモーションが高精度に推定されない場合、式 (12) のモデル化誤差となり、復元画像に偽色が発生することが確認されているためである。入力の RAW データに強いエイリアシングが入るほど超解像の効果が高まる反面、マッチング精度が低下することが考えられ、そのような条件下でも高精度な “RAW-to-RAW registration”<sup>(15)</sup> を行うことが重要な課題となる。

### 3.3 画像復元アルゴリズムの概要

提案手法による高解像度カラー画像の合成は、次のようなアルゴリズムによって実現できる。

- (1) 入力 RAW 画像列  $y_1, \dots, y_M$  を取得する。
- (2) CFA パターンが未知の場合は、感度配列  $m_c(i_1, i_2)$  のキャリブレーションを行う。R, G, B の各フィルタに対応する画素の位置は、それぞれ赤、緑、青の物体を撮影したときに最も輝度の大きい画素の位置に一致することから推定することができる。
- (3) 入力 RAW 画像列を線形 demosaicking によりカラー化する。得られたカラー画像の輝度成分に対して、画像の位置ずれ  $d_k = [d_{xk}, d_{yk}]^T$  を推定する。

- (4) 所望の解像度倍率  $f$  を設定する。

次に、式 (13) の最適化を行う。最適化手法として最急降下法を用いると次のようなアルゴリズムとなる。

- (5)  $n = 0$  とする。基準となる入力画像 1 枚から線形補間により高解像度カラー画像を生成する。この画像を復元画像の初期値  $z^{(0)}$  とする。
- (6) 画像  $z^{(n)}$  の輝度成分のエッジ方向解析<sup>(8)</sup> より、正則化項  $f_2(z_Y)$  の重み  $\Lambda_d$  を計算する。
- (7) 次式により画像  $z^{(n)}$  を更新する。

$$z^{(n+1)} = z^{(n)} - \alpha \left[ \frac{\partial}{\partial z} \sum_{m=1}^3 f_m(z) \right]_{z=z^{(n)}} \quad (22)$$

$$\left. \frac{\partial f_1(z)}{\partial z} \right|_{z=z^{(n)}} = \sum_{k=1}^M A_k^T (A_k z^{(n)} - y_k) \quad (23)$$

$$\left. \frac{\partial f_2(z_Y)}{\partial z} \right|_{z=z^{(n)}} = \sum_{d \in D} T_Y^T P_d^T \Lambda_d^T \Lambda_d P_d T_Y z^{(n)} \quad (24)$$

$$\left. \frac{\partial f_3(z_{Cb}, z_{Cr})}{\partial z} \right|_{z=z^{(n)}} = \lambda_c \sum_{C \in \{Cb, Cr\}} T_C^T H^T H T_C z^{(n)} \quad (25)$$

ここで、 $\alpha$  は最急降下法のパラメータ (ステップサイズ) である。 $T_T, T_{Cb}, T_{Cr}$  はそれぞれ RGB 信号から  $Y, Cb, Cr$  成分を求める演算を表す行列である。

- (8)  $n = n + 1$  と置き、(6) に戻る。
- (9) 繰返し計算の終了条件としては、繰返し計算の収束を判定するか、またはあらかじめ定めた繰返し回数で打ち切る。

## 4. 実験結果

ここでは、前章のアルゴリズムを用い、提案手法の有効性を検証する実験を行った。

### 4.1 シミュレーションによる評価

図 7(a) はシミュレーションで用いた高解像度参照画像であり、図 7(b) は参照画像 (a) から単板撮像素子による撮像モデルを用いて合成した画像である。なお、ここでは撮像モデルにおけるモーション  $s_k$  は既知とした。

図 7(c) は 1 枚の入力画像を線形 demosaicking した結果である。(d) は (c) を線形補間により高解像度化した結果である。(e) は 8 枚の入力画像のそれぞれに線形 demosaicking を施し、得られたカラー画像列の各色チャンネルに超解像<sup>5)</sup>を適用した結果である(図 1 参照)。(f) は 1 枚の入力画像に既存の demosaicking 手法<sup>2)</sup>を施した結果を示す。(g) は (f) をバイキュービック補間により高解像度化した結果である。(h) は 8 枚の入力画像のそれぞれに demosaicking<sup>2)</sup>を施し、得られたカラー画像列の各色チャンネルに超解像<sup>5)</sup>を適用した結果である。(d)(e)(g)(h) の画像拡大倍率は水平・垂直方向ともに 2 倍である。(d)(g) は比較的単純な処理により RAW データから高解像度カラー画像を生成したものであり、典型的なアーチファクト(ボケ、偽色)が発生していることが分かる。また、(e)(h) では 2 段階の処理により解像感の改善が見られるが、第 1 段階の処理で生じたアーチファクトが最終結果にも影響していることが分かる。

図 7(i)(j)(k) に種々の条件での提案手法による画像復元結果を示す。図 7(k) は 8 枚の入力画像から解像度が 2 倍のカラー画像を復元した結果である。(e)(h) と (k) は同じ 8 枚の画像列から解像度が 2 倍の画像を復元したものであるが、提案手法 (k) では、高い空間周波数成分も忠実に復元されており、さらに偽色の発生も抑制されていることが確認される。本手法の実際の適用場面では、必ずしも複数の入力画像が得られるとは限らない。この場合も、提案手法の特殊なケースとして、1 枚の入力画像から同一のアルゴリズムによって画像復元を行うことができる。図 7(i)(j) は 1 枚の入力画像からそれぞれ解像度が 1 倍および 2 倍の画像を生成した結果である。

図 8(a)(b)(c) では 8 枚の入力画像列の位置合わせをし、すべての画像データを重ねたうえで、補間により高解像度の画素値を求めた結果である。(a) は最近傍点補間、(b) は三角形分割による線形補間、(c) は三角形分割によるキュービック補間を用いている。同

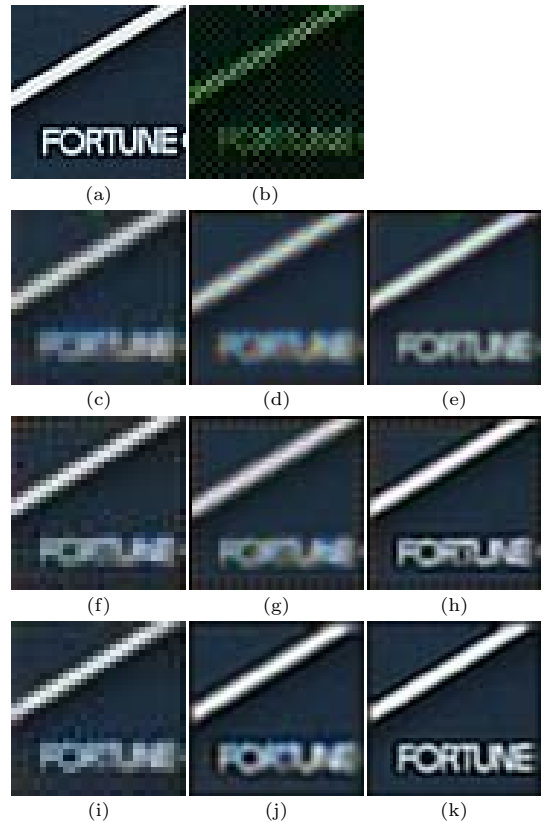


図 7 復元画像。(a) 参照画像。(b) 入力画像。(c) 線形 demosaicking。(d) 線形 demosaicking + 線形補間。(e) 線形 demosaicking + グレースケール超解像<sup>5)</sup>。(f) Demosaicking<sup>2)</sup>。(g) Demosaicking<sup>2)</sup> + バイキュービック補間。(h) Demosaicking<sup>2)</sup> + グレースケール超解像<sup>5)</sup>。(i) 提案手法,  $f = 1, M = 1$ 。(j) 提案手法,  $f = 2, M = 1$ 。(k) 提案手法,  $f = 2, M = 8$

Fig. 7 Reconstructed images. (a) Reference. (b) Input. (c) Linear demosaicking. (d) Linear demosaicking and interpolation. (e) Linear demosaicking and conventional super-resolution<sup>5)</sup>. (f) Demosaicking<sup>2)</sup>. (g) Demosaicking<sup>2)</sup> and bi-cubic interpolation. (h) Demosaicking<sup>2)</sup> and conventional super-resolution<sup>5)</sup>. (i) Proposed,  $f = 1, M = 1$ . (j) Proposed,  $f = 2, M = 1$ . (k) Proposed,  $f = 2, M = 8$ .

じ入力画像列から提案手法により復元した画像 (d) と (a)(b)(c) を比較しても、(d) においてアーチファクトが低減されていることが分かる。

本シミュレーション実験では、提案手法の定量的評価として、高解像度参照画像と復元画像の誤差を二乗平均平方根 (Root Mean Square: RMS) により算出した。図 9 は提案手法において、入力画像の枚数と RMS 値の関係を示す。同図において、図 1 のような 2 段階の処理 (複数の入力画像のそれぞれに demosaicking (線形補間、あるいは文献 2) の手法) を施し、得られたカラー画像セットに超解像<sup>5)</sup>を適用) を施した



図 8 復元画像 . (a) 最近傍点補間 (複数画像) . (b) 線形補間 (複数画像) . (c) キュービック補間 (複数画像) . (d) 提案手法,  $f = 2, M = 8$

Fig. 8 Reconstructed images. (a) Nearest neighbor interpolation (multi-frame). (b) Linear interpolation (multi-frame). (c) Cubic interpolation (multi-frame). (d) Proposed,  $f = 2, M = 8$ .

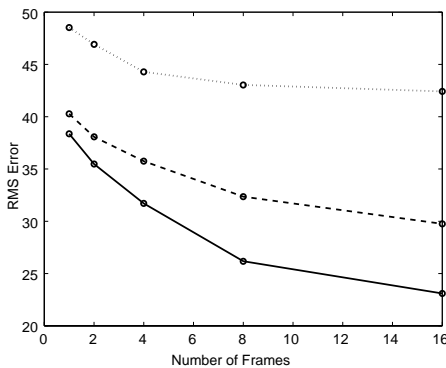


図 9 RMS 誤差 . 点線 : 線形補間 + 超解像<sup>5)</sup>, 破線 : demosaicking<sup>2)</sup> + 超解像<sup>5)</sup> 実線 : 提案手法

Fig. 9 RMS error. Dotted line: linear interpolation, dashed line: demosaicking<sup>2)</sup> and bi-cubic interpolation, solid line: proposed.

結果の RMS 値も示されている . 特に複数枚の入力画像が得られる場合 , 提案手法による画像復元の効果が顕著に表れることが分かる .

次に , 入力画像間の位置ずれを未知として推定した場合に , その推定誤差が復元画像に及ぼす影響について実験した . ここでは , 前述の実験の図 7(k) と同様に , 8 枚の画像から解像度が水平方向・垂直方向ともに 2 倍の画像を復元する実験を行った . ただし , ここでは入力画像 8 枚のうち 1 枚は水平方向の位置ずれに誤差が含まれるとし , その大きさが ( 高解像度画像座標において ) 0, 0.25, 0.5, 0.75, 1 [画素] の場合の復元画像を図 10 に示す . この結果から , 0.25 画素程度の位置ずれ推定誤差であれば , 画質が大きく損なわれ

ないことが分かる .

本実験ではまた , 提案手法の計算時間を Pentium 3 933 MHz の PC で計測した . 入力画像サイズが  $30 \times 30$  , 解像度倍率が水平・垂直方向ともに 2 倍の条件で CPU 時間を計測すると次のようになる . 提案手法 ( 入力 8 枚 ) : 2.10 sec , 提案手法 ( 入力 1 枚 ) : 0.79 sec , Demosaicking<sup>2)</sup> + バイキュービック補間 ( 入力 1 枚 ) : 0.31 sec , 線形補間 ( 入力 1 枚 ) : 0.13 sec . なお , 提案手法の計算時間は , モーション推定やフレーム選び等の前処理を除いた処理の「繰返し計算 1 回あたりの秒数」を示す . 提案手法での繰返し計算の回数は通常数十回程度であり , 入力画像の枚数が多いほど繰返し回数が少なくてすむ傾向がある .

#### 4.2 文書画像による実験

ここでは , 実際に単板カメラを用いて , 手持ちで撮影した動画から高解像度画像を復元する実験を行った . 使用したカメラは , Point Grey Research 社製の Dragonfly であり , CFA は既知 ( Bayer パターン ) である . 手ぶれによる微小な平行移動を考慮し , 3.2 節に示した位置ずれ推定を行った . 図 11 (a) は入力画像の 1 枚目である . 復元する高解像度画像は撮像素子の物理的解像度に対し , 水平方向 , 垂直方向それぞれ 4 倍の解像度を持つ . 図 11 (b)(c) は既存の手法 ( 線形補間 , および demosaicking<sup>2)</sup> ) , 図 11 (d) は 64 枚の画像を用いた提案手法による処理結果を示す . 提案手法が格段に優れた復元画像をもたらすことが分かる .

#### 4.3 自然画像による実験

ここでは , 種々の自然画像を対象とした実験結果を示す . 撮影にはソニー製 3CCD カムコーダを使用し , そのカラー出力画像をもとに Bayer フィルタをかけた RAW データを生成した . 図 12 (1) , (2) はカメラの手ぶれを利用して静止した対象を撮影した結果であり . 図 12 (3) はカメラを固定し , 風で揺れている対象を撮影した結果である . 後者の場合は , 画像全体が異なるモーションで表せないため , 画像内の局所的なブロックに対してモーションパラメータを推定している .

実画像を用いた実験では , より多くの入力画像を用いるほど画像の復元精度が高くなるとは限らない . その原因としては , 入力画像間のモーションに推定誤差があること , モーションが 2 パラメータの並進運動では正確に表現できないこと , 時間の経過とともに明るさが変化すること等が考えられる . そのため , 本実験では , 利用可能な入力画像列の中から , 次の基準で入力画像を選択している .

- 入力の基準画像から時間的に近い画像を優先的に選択する .



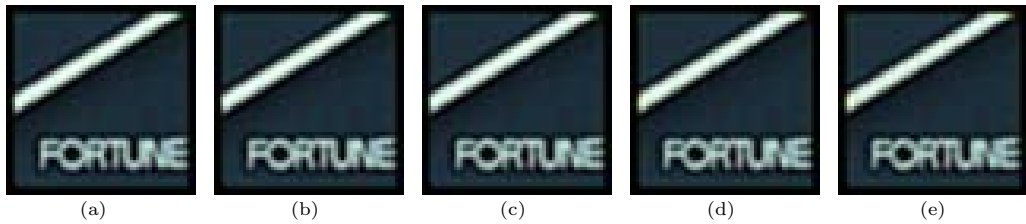


図 10 位置ずれ推定誤差の復元画像への影響 . (a) 0 画素 , (b) 0.25 画素 , (c) 0.5 画素 , (d) 0.75 画素 , (e) 1 画素

Fig. 10 Motion estimation error affecting the image estimate. (a) 0 pixels, (b) 0.25 pixels, (c) 0.5 pixels, (d) 0.75 pixels, (e) 1 pixels.

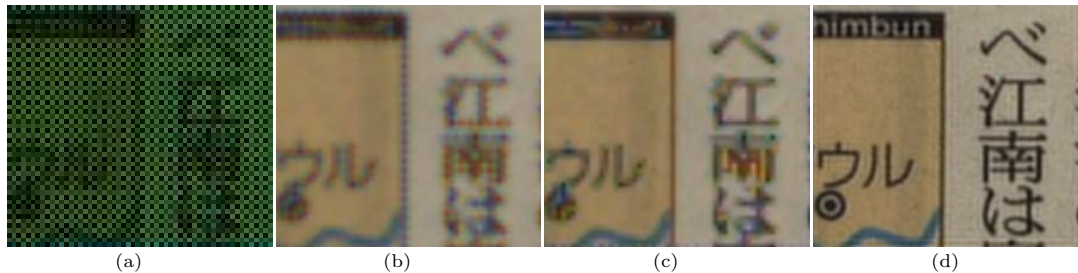


図 11 復元画像 . (a) 入力画像 . (b) 線形 demosaicking + 線形補間 . (c)

Demosaicking<sup>2</sup>) + バイキュービック補間 . (d) 提案手法 ,  $f = 4$  ,  $M = 64$

Fig. 11 Reconstructed high-resolution images. (a) Observed color mosaic. (b) Linear demosaicking and interpolation. (c) Demosaicking<sup>2</sup>) and bi-cubic interpolation. (d) Proposed,  $f = 4$  ,  $M = 64$ .

- $\text{mod}(d_1, 2), \text{mod}(d_2, 2), \dots, \text{mod}(d_M, 2)$  が領域  $[0, 2) \times [0, 2)$  内で均一に分布するような入力画像を優先的に選択する .
- 入力画像間の位置ずれの推定における画像の一致度評価値 (SSD, SAD 等) が低い画像は選択しない .

復元する高解像度画像は撮像素子の物理的解像度に対し, 図 12(1)(2) では水平, 垂直方向ともに 4 倍, 図 12(3) では 2 倍の解像度を持つ . 図 12(b) は線形補間, (c) は demosaicking<sup>2</sup>) + バイキュービック補間, (d) は提案手法による処理結果 (入力画像の枚数は (1)(2) が 64 枚, (3) が 16 枚である) を示す . 提案手法が本実験でも効果的であることが確認される .

## 5. おわりに

本研究では, 単板撮像素子の RAW データと所望の解像度のカラー画像の関係を直接的なモデルで定式化し, 高解像度カラー画像を生成するアルゴリズムを提案した . 提案手法は撮像系の製造段階で決められた物理的分解能を超える画像を事後的にソフトウェアで生成することができるという利点も持つ . また, 入力画像の枚数や解像度拡大率が任意に設定できる . 適用可

能な CFA パターンが Bayer パターン等に特化したものではないという点でも一般性のある手法である .

近年, RAW データが取得可能なカメラが普及しつつあることから, 今後提案手法が適用できる場面が広がっていくことが予想される .

## 参 考 文 献

- 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, and Prescott, M.A.: Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients, United States Patent 5,373,322 (1994).
- 3) Huang, T.S. and Tsay, R.Y.: Multiple frame image restoration and registration, *Advances in Computer Vision and Image Processing*, Vol.1, Huang, T.S. (Ed)., pp.317-339, JAI Press Inc, Greenwich (1984).
- 4) Irani, M. and Peleg, S.: Improving resolution by Image Registration, *CVGIP: Graph. Models Image Process.*, Vol.53, pp.231-239 (Mar. 1991).

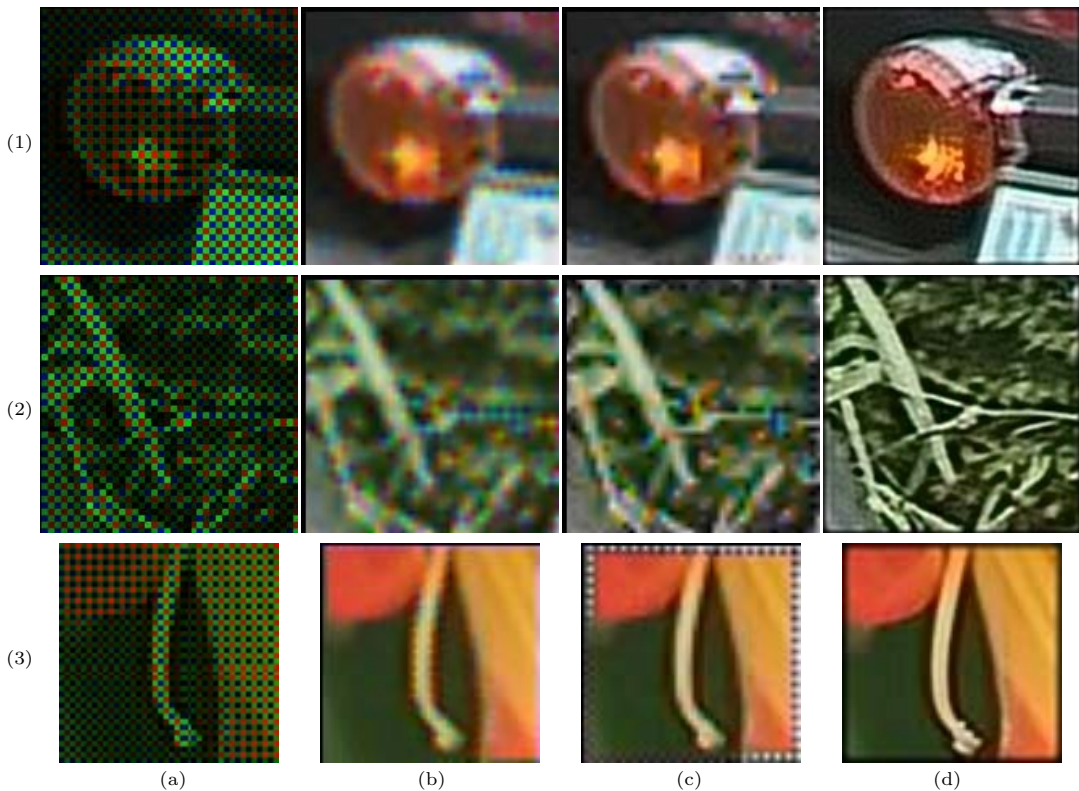


図 12 復元画像 . (a) 入力画像 . (b) 線形 demosaicking + 線形補間 . (c) Demosaicking<sup>2</sup> + バイキュービック補間 . (d) 提案手法 ( (1):  $f = 4, M = 64$ , (2):  $f = 4, M = 64$ , (3):  $f = 2, M = 16$  )

Fig. 12 Reconstructed high-resolution images. (a) Observed color mosaic. (b) Linear demosaicking and interpolation. (c) Demosaicking<sup>2</sup> and bi-cubic interpolation. (d) Proposed ((1):  $f = 4, M = 64$ , (2):  $f = 4, M = 64$ , (3):  $f = 2, M = 16$ ).

- 5) Hardie, R.C., Barnard, K.J. and Armstrong, E.E.: Joint MAP Registration and High-Resolution Image Estimation using a Sequence of Undersampled Images, *IEEE Trans. Image Processing*, Vol.6, pp.1621–1633 (1997).
- 6) Tekalp, A.M., Ozkan, M.K. and Sezan, M.I.: High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration, *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, San Francisco, CA., Vol.III, pp.169–172 (Mar. 1992).
- 7) Schultz, R.R. and Stevenson, R.L.: Improved definition video frame enhancement, *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Detroit, MI., Vol.IV, pp.2169–2172 (May 1995).
- 8) Shin, J., Paik, J., Price, J.R. and Abidi, M.A.: Adaptive regularized image interpolation using data fusion and steerable constraints, *SPIE Visual Communications and Image Processing*, Vol.4310 (Jan. 2001).
- 9) Capel, D. and Zisserman, A.: Automated Mosaicing with Super-resolution Zoom, *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (1998).
- 10) Bayer, B.E.: Color Imaging Array, United States Patent 3,971,065 (1976).
- 11) Shimizu, M. and Okutomi, M.: Precise sub-pixel estimation on area-based matching, *Proc. 8th International Conference on Computer Vision*, pp.90–97 (Jul. 2001).
- 12) Shimizu, M. and Okutomi, M.: An Analysis of Sub-Pixel Estimation Error on Area-Based Image Matching, *Proc. 14th International Conference on Digital Signal Processing (DSP2002)*, Vol.II, pp.1239–1242 (W3B.4) (Jul. 2002).
- 13) Shimizu, M. and Okutomi, M.: Two-Dimensional Simultaneous Sub-Pixel Estimation on Area-Based Image Matching, *Proc. Asian Con-*

*ference on Computer Vision (ACCV2004)*, pp.854-859 (P-93) (Jan. 2004).

- 14) Gotoh, T. and Okutomi, M.: Color Super Resolution from a Single-CCD, *CD-ROM Proc. IEEE Workshop on Color and Photometric Method in Computer Vision (CPMCV, in conjunction with ICCV)* (Oct. 2003).
- 15) Gotoh, T. and Okutomi, M.: Direct Super-Resolution and Registration Using Raw CFA Images, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR2004)* (Jun. 2004).

(平成 16 年 1 月 15 日受付)

(平成 16 年 3 月 4 日採録)

(担当編集委員 八木 康史)



後藤 知将

2001 年東京工業大学工学部制御システム工学科卒業。2003 年東京工業大学大学院情報理工学研究科情報環境学専攻修士課程修了。同年(財)理工学振興会(東工大 TLO)プロジェクト研究員。コンピュータビジョン, 信号処理の研究に従事。同年ソニー(株)入社。



奥富 正敏(正会員)

1981 年東京大学工学部計数工学科卒業。1983 年東京工業大学大学院理工学研究科制御工学専攻修士課程修了。同年キヤノン(株)入社。1987 年~1990 年カーネギーメロン大学コンピュータサイエンス学科客員研究員。1994 年東京工業大学大学院情報理工学研究科情報環境学専攻助教授。2002 年同大学院理工学研究科機械制御システム専攻教授。コンピュータビジョン, 画像処理, 画像計測に関する研究に従事。工学博士。計測自動制御学会, 電子情報通信学会, 日本ロボット学会, 画像電子学会, IEEE 各会員。