

# カメラの絞りに色フィルタを用いた奥行き推定と前景マット抽出 Depth Estimation and Foreground Matting using Color Filters in a Camera Aperture

坂東 洋介<sup>†,‡</sup>

Yosuke Bando

<sup>‡</sup>(株)東芝

TOSHIBA Corporation

西田 友是<sup>†</sup>

Tomoyuki Nishita

<sup>†</sup>東京大学

The University of Tokyo

## 1. はじめに

近年カメラを加工して、撮影した画像からシーンの奥行きなど、付加的な情報を得る手法の開発が盛んである。しかしながら多くの手法は複雑で精密な加工が必要で機器が高価になりやすい。また、校正が必要となることも多く、設置に時間がかかり持ち運びが容易でない。

本論文は、カメラレンズの絞り位置に RGB 三色のフィルタを配置してシーンを撮影することにより、一枚の画像からシーンの奥行きと前景領域(マットと呼ぶ)を自動抽出する手法を提案する。レンズ開口部を RGB それぞれの波長帯のみを通す三領域に分割することで、RGB 画像間に奥行きに応じた視差(色ずれ)が生じることを用いて奥行きを推定する。さらに本論文は、前景物体に焦点を合わせることで背景の色がずれて撮影されることを利用して、前景領域を抽出する。前景領域と背景領域の境界部において、色のずれ量を考慮して前景色と背景色の混合率を求める。これにより毛のように複雑で不明確な輪郭を持つ前景に対しても高精度なマットを自動抽出できる。

本手法で得られた奥行きと前景マットの利用例として、視点の変更、焦点ボケの除去、異なる背景への合成を含む画像編集例を示す。

## 2. 関連研究

### 2.1. 色フィルタを使用した奥行き推定

我々の知る限り、カメラの絞りに色フィルタを配置して奥行きを推定する手法は過去に二例提案されている。Amariら[1]は波長帯の異なる画像間のマッチングをとるために画像にハイパスフィルタをかけてから局所ウィンドウ内の輝度差を最小化する指標を用いたが、この指標は著者らも議論している通り、画像間の輝度差を補償するには不十分であった。Changら[5]は局所ウィンドウ内で輝度の総和を一致させるスケールリングをしてから輝度差を最小化する指標を用いたが、フラッシュで撮影対象物体に斑点模様を投射して強いエッジを作る必要があった。そのため、取得した奥行き画像にテクスチャを貼るためにはフラッシュをたかずにもう一枚撮影する必要があった。我々は能動的な照明

なしで撮影した画像のみから奥行きを推定するためのよりよいマッチング指標を提案する。また、色フィルタを用いた前景マット抽出は我々が初めて提案する。

### 2.2. その他のカメラ加工による奥行き推定

Adelsonら[2]は、撮像面にレンズアレイを敷くことでカメラに入射する光を方向別に記録できることを示し、単眼で撮影した多視点画像による奥行き推定を行った。Ngら[18]はこの考えを手持ちカメラで実現したほか、多視点画像がカメラ内の light field に対応することを用いて、効率よく焦点ボケを変更した画像を生成する手法を提案した[19]。また、撮像面側でなくカメラのレンズ側(開口部)で光路を分割して同様のことを実現する手法が提案された[8,15]。我々の手法もこの分類に該当するが、提案法は追加の装置が色フィルタのみで簡素であり、かつ一度の露光で撮影できることが差異となる。

視差によらず焦点ボケにより奥行きを推定する方法として、レンズ開口部を同心円状に分割して異なるボケ方をした複数の画像を得る方法[10]、画像処理で除去可能な程度の光の点をプロジェクタで投影する方法[17]、開口部に模様をついたマスクを置く方法[13,20]が提案されている。開口部にマスクを置く手法は加工を最小限に留めたカメラによるシーンの奥行き推定を可能にした点で我々の手法に近い(ただしVeeraraghavanらはマスクを撮像面に置いてレンズアレイなしで light field を取得する手法をも提案している)。我々の手法は視差によるため、ボケ推定では難しい、焦点の合った位置より遠いか近いかの判定が容易であることと、ボケ除去が不要なため計算が高速であることが利点である。

### 2.3. 前景マット抽出

一枚の画像からマットを抽出する手法の多くは、ユーザに画像を「確実に前景」「確実に背景」「不明」の三領域に塗り分けさせ(トライマップと呼ぶ)、「確実に前景」「確実に背景」領域の色分布を利用して「不明」領域のマットを計算する[6,12,21]。

一方で、マット抽出を自動化する手法は、複数台の

カメラとそれらの位置合わせを必要とし、設置に時間がかかり簡単には移動できない[11,16]。提案法は一台のカメラで一度露光するだけであり、持ち運びが容易で位置合わせが不要である。

### 3. 色フィルタ付きレンズ

図 1(a)は本論文で使用する色フィルタの配置であり、これを絞り位置に装着したレンズを図 1(b)に示す。R, G, B の各領域の形状が合同で、かつ開口部の中心から各領域の中心への変位（図 1(a)の矢印）が撮像面の  $x$  軸と  $y$  軸に沿うように領域を配置した。これにより、焦点の合った奥行きより遠い点は R が右、G が上、B が左方向にそれぞれずれて観測される。焦点の合った奥行きより近い点ではずれ方向が逆になる。この原理を図 2 に R と G の二色分割の場合で示した。焦点の合っている奥行きでは画像はずれない（図 2 左）が、焦点の合っていない奥行きではシーンの同じ点が色ごとにずれて観測される（図 2 右）。

我々はプロトタイプとしてカメラに Canon EOS40D、レンズに Canon EF50mm f/1.8 II、フィルタの RGB にそれぞれ Fujifilter SC-58, BPB-53, BPB-45 を使用した。ただしカメラ本体には加工はしていない。フィルタとカメラの色分解特性は必ずしも一致しないので、図 2 のように R のフィルタを通った光だけが画像の R 成分として記録されるわけではなく多少の相互作用がありうる。我々は[1]と同様の手法でこの相互作用を計測し、色変換を行いこれを低減した。詳細は付録 A.1 参照。

図 1(b)に示したプロトタイプのレンズのボケ関数を図 3 に示す。これは焦点の合った奥行きより遠くに置いた点光源を暗室で観測したときの像であり、上記色変換をかけた後のものである。R と G の左右の角がわずかにレンズ筐体に遮蔽されて欠けている以外は、色フィルタの菱形の形状がほぼそのまま観測されている。レンズの外から見た図 1 とは配置が左右逆になっていることに注意。色変換によりほぼ完全に各フィルタを通った光が分離できていることが分かる。ただし、色変換前でも相互作用はそれほど大きくなかった。

フィルタの配置を撮像面の軸に一致させるのは目視と手動による調整で十分であった。カメラの向きによらず色ずれ方向が  $x$  軸と  $y$  軸に沿うので、以降の画像処理が容易になる上、カメラを移動させるたびに較正が必要になることもない。図 4 に撮影した画像の例と、それを R, G, B 成分に分離した画像を示す。

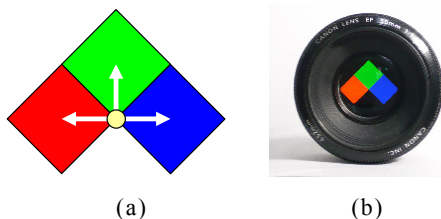


図 1: (a)色フィルタの配置。(b)レンズへの装着例。

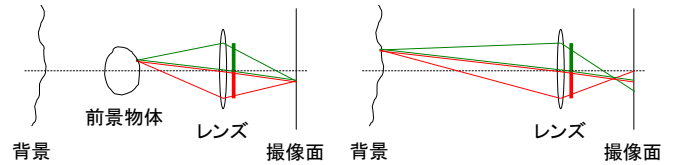


図 2: 色フィルタによる視差の原理

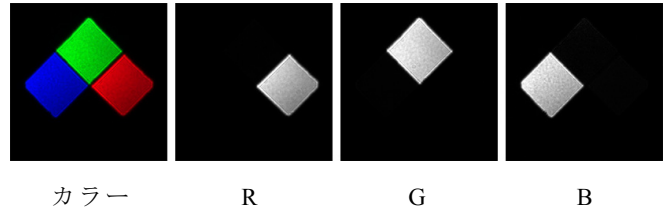


図 3: 色変換後のボケ関数



図 4: 色フィルタ付きレンズで撮影した画像と、RGB 別の表示。ずれが分かりやすいように白線を記入してある。焦点を合わせた犬のぬいぐるみにはずれがない。

## 4. 画像処理

### 4.1. 奥行き推定

撮影した画像の RGB 成分  $I_r, I_g, I_b$  は三視点のステレオ画像となる。すると、仮想的な中央の視点をリファレンス座標にとれば、3 節で述べたようにフィルタの配置に従い R が右、G が上、B が左方向にずれる。そこで、リファレンス座標  $(x, y)$  における視差を  $d$  画素とすると、 $I_r(x+d, y), I_g(x, y-d), I_b(x-d, y)$  が対応するから、これらが最もよく一致する  $d$  を視差の推定値とすることができる。しかし、観測波長帯が異なるため対応点の輝度レベルは一致しないので、単純な輝度の差分をマッチング指標にすることはできない。

そこで、提案法は自然画像の色成分は局所的に線形関係を示す傾向が高いことを利用する[9,12]。例として、通常のカメラで撮影した図 5(a)に示す画像の一部の色分布を RGB の三次元空間にプロットすると図 5(b)左上のようになる。よって、 $(x, y)$  の周りに局所ウィンドウ  $w(x, y)$  を考え、そこに含まれる画素色の集合を  $S(x, y; d) = \{(I_r(s+d, t), I_g(s, t-d), I_b(s-d, t)) \mid (s, t) \in w(x, y)\}$  とすると、真の  $d$  のときに分布が最も直線に近くなる

と考えられる。実際、図 5(a)の画像では  $d = 0$  が真の値であるが、 $d$  を変化させながら色分布をプロットすると図 5(b)に示したように直線性が急速に崩れてゆく。

本論文では RGB 間の線形関係を測る指標として以下を用いる。

$$L(x, y; d) = \lambda_0 \lambda_1 \lambda_2 / \sigma_r^2 \sigma_g^2 \sigma_b^2 \quad (1)$$

ただし  $\lambda_0, \lambda_1, \lambda_2$  ( $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq 0$ ) は色分布  $S(x, y; d)$  の主成分軸に沿った分散 (すなわち  $S(x, y; d)$  の共分散行列  $\Sigma$  の固有値)、 $\sigma_r^2, \sigma_g^2, \sigma_b^2$  は色分布の R, G, B 軸に沿った分散である。ここでは表記の簡単のため右辺の  $x, y, d$  への依存については省略した。

行列の性質から  $\lambda_0 + \lambda_1 + \lambda_2 = \sigma_r^2 + \sigma_g^2 + \sigma_b^2$  (=  $\text{trace}(\Sigma)$ ) であるので、固有値間の大小に差があるほど式(1)は小さくなる。直線的な分布 ( $\lambda_0 \gg \lambda_1, \lambda_2$ ) であれば  $L$  は 0 に近く、完全無相関 (固有値が R, G, B 軸に沿った分散に等しい) のとき最大値  $L = 1$  をとる。これは二画像間のステレオマッチングに使用される指標の 1 つである Normalized Cross Correlation (NCC)[14] の拡張ともみなすことができる。詳細は付録 A.2 参照。

画像の各点、および取りうる範囲の各視差  $d$  について  $L(x, y; d)$  を計算する必要があるが、Summed Area Table[7]を用いて  $I_r, I_g, I_b$  とそれらの二乗と積の表を計算しておけば、局所ウィンドウのサイズに依存せずに共分散行列  $\Sigma$  の各要素を計算することができる。また、式(1)の分子は行列の性質から  $\lambda_0 \lambda_1 \lambda_2 = \det(\Sigma)$  であるので固有値を実際に求める必要はない。

図 4 の画像に対して  $L(x, y; d)$  を計算して、各画素で最小値を与える  $d$  をとった結果を図 6(a)に示す。各画素で独立に推定するとノイズに弱いので、グラフカット法[4]により近傍画素間の推定値のスムーズさも考慮して最小化を行うと図 6(b)の結果を得た。

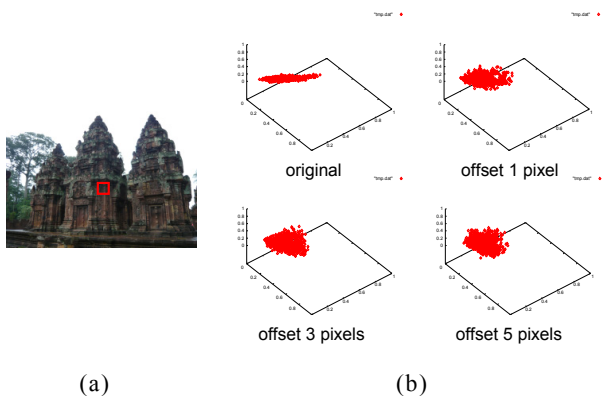


図 5: 局所ウィンドウ内の色の RGB 空間中の分布例。(a) 色サンプルをとった画像。局所ウィンドウを赤矩形で示した。(b) RGB 空間中の色分布。グラフの 3 軸はそれぞれ R, G, B の輝度レベルを示す。左上が元の画像における色分布、その他は図中に記載した視差を仮定して画像を色ずれさせてから色分布を示した。

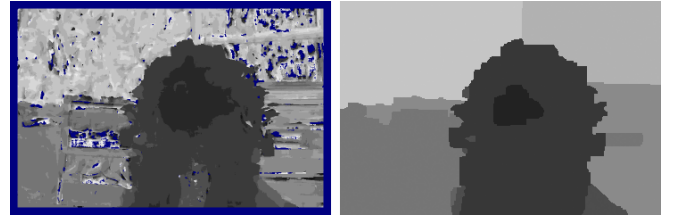


図 6: 図 4 の画像に対する色ずれ量の推定値で、輝度が高いほど色ずれ量が多い (焦点を合わせた距離より遠い) ことを表す。(a)局所的な推定値。青い領域は画面端であるか、周囲と極度に値が異なるため信頼できる推定値がないことを示す。(b)グラフカット法による最適化結果。

## 4.2. 前景マツト抽出

画像から前景物体を切り出すマツト抽出は、物体を異なる背景に合成するという、画像合成の基礎的な処理に不可欠な技術である。マツトとは、式(2)に示すように、画像  $I$  が前景色  $F$  と背景色  $B$  の線形混合であるとするモデルにおける混合率  $\alpha$  ( $\in [0, 1]$ ) を各画素  $(x, y)$  において保持するグレースケール画像である。

$$\begin{aligned} I_r(x, y) &= \alpha(x, y)F_r(x, y) + (1 - \alpha(x, y))B_r(x, y) \\ I_g(x, y) &= \alpha(x, y)F_g(x, y) + (1 - \alpha(x, y))B_g(x, y) \\ I_b(x, y) &= \alpha(x, y)F_b(x, y) + (1 - \alpha(x, y))B_b(x, y) \end{aligned} \quad (2)$$

式(2)は画素数  $N$  に対しマツトに関する制約式を  $3N$  個与える。よって、一枚の画像  $I$  が与えられただけでは、未知数が  $\alpha, F_r, F_g, F_b, B_r, B_g, B_b$  の  $7N$  個あるため、このままではマツトの抽出は困難である。

そこで我々は、提案カメラレンズにより前景に焦点を合わせて撮影すると主に背景  $B$  のみはずれることを利用して解を制約する。より正確には、 $\alpha$  値が中間的な値をとる、前景色と背景色の混合領域がレンズの焦点深度内に入っていれば R, G, B 画像で  $\alpha(x, y)$  に視差はないとみなせるので、式(2)が有効となる。これは[11]の仮定と同様である。

まず、4.1 節で求めた視差  $d(x, y)$  を利用し、閾値と比較して画像を前景領域と背景領域に二分割したのち、二領域の境界部分を広げて「不明」領域とし、トライマップを作る (図 7(a))。次に、「不明」領域に「確実に前景」領域と「確実に背景」領域から視差  $d(x, y)$  を伝播する (図 7(b)(c))。これにより、「不明」領域において前景視差  $d_F(x, y)$  と背景視差  $d_B(x, y)$  が得られる。

この 2 つの視差を用いて、以下に示す前景色ずれ度  $M_F(x, y)$  と背景色ずれ度  $M_B(x, y)$  を定義する。

$$\begin{aligned} M_F(x, y) &= e_F(x, y, d_F(x, y)) - e_F(x, y, d_B(x, y)) \\ M_B(x, y) &= e_B(x, y, d_B(x, y)) - e_B(x, y, d_F(x, y)) \end{aligned} \quad (3)$$

ただし  $e_F(x, y; d), e_B(x, y; d)$  は式(4)に示す色ずれ誤差

をそれぞれ前景色  $F$ 、背景色  $B$  について計算した値である。

$$e(x, y; d) = \sigma_r^2 + \sigma_g^2 + \sigma_b^2 - \lambda_0 \quad (4)$$

式(4)は 4.1 節における局所ウィンドウ内の色分布に当てはめた直線（主成分軸）からの距離の二乗平均であり（導出は付録 A.3 参照）、線形色モデルからの誤差を表す。ここで式(1)と異なる指標を用いる理由は、第一に、式(1)は RGB のうち一成分のみが変化する領域（RGB 軸に平行な直線状の分布）においても値が大きくなるからである。これは平坦な領域とのマッチングを好む傾向を避ける意味で、奥行き推定では有効であったが、色ずれ誤差の指標としては望ましくない。第二に、値の範囲が正規化された式(1)よりも、直接的な誤差を考えた方が、誤差が大きいところほど以下で説明する最小化計算の影響が大きくなる。

式(3)の前景色ずれ度  $M_F(x, y)$  の意味について説明する。前景色は前景色ずれ量  $d_F(x, y)$  を適用して画像の色ずれを相殺したときに線形色モデル誤差  $e_F(x, y; d_F(x, y))$  が小さくなるのが期待される。逆に、背景色ずれ量  $d_B(x, y)$  を適用して画像の色ずれを相殺すると、誤った色ずれ量で復元するために色ずれは修正されず、線形色モデル誤差  $e_F(x, y; d_B(x, y))$  が大きくなるのが期待される。よって  $M_F$  は小さくなるのが望ましい。背景色ずれ度  $M_B$  も同様である。そこで我々は、 $M_F$  と  $M_B$  を最小化するような解を求める。最小化すべきエネルギー関数  $E(\alpha, F, B)$  は以下ようになる。

$$E(\alpha, F, B) = E_m(\alpha, F, B) + \kappa_\alpha E_\alpha(\alpha) + \kappa_s E_s(F) + \kappa_s E_s(B) + \kappa_c \sum_{(x,y)} M_F(x, y) + \kappa_c \sum_{(x,y)} M_B(x, y) \quad (5)$$

ただし  $E_m = \sum_{(x,y)} |1 - \alpha F - (1 - \alpha)B|^2$  は式(2)が成立することを要求する項、 $E_\alpha$ ,  $E_s$  はそれぞれ  $\alpha$ ,  $F$ ,  $B$  のスムーズさの制約で、 $E_s$  には単純な近傍画素との差の二乗和を使用するが、 $E_\alpha$  については精度のよいマットを生成することで知られる [12] の方法を使用する。 $\kappa_\alpha$ ,  $\kappa_s$ ,  $\kappa_c$  はそれぞれの項の影響を調整する係数である。

色ずれ量最小化という条件を加えたにせよ、 $M_F$  と  $M_B$  のため問題は非線形である上に、観測が画像一枚であることに変わりはないので、式(5)の  $7N$  個の未知数を一度に最適化するのは現実的でない。そこで我々は第一に得たい対象であるマット  $\alpha$  の  $N$  個の未知数について最適化することを考え、 $F$ ,  $B$  は、 $\alpha$  が与えられれば  $E_m(\alpha, F, B) + \kappa_s E_s(F) + \kappa_s E_s(B)$  を最小化することで得られる  $\alpha$  の従属変数  $F(\alpha)$ ,  $B(\alpha)$  と考えてエネルギー関数を  $\alpha$  のみの関数  $E'(\alpha)$  とする。

$$E'(\alpha) = E_\alpha + \kappa'_c \sum_{(x,y)} M_{F(\alpha)} + \kappa'_c \sum_{(x,y)} M_{B(\alpha)} \quad (6)$$

ただし  $\kappa'_c = \kappa_c / \kappa_\alpha$  である。次に、第二、第三項を現在の  $\alpha$  の見積もりにおいて二次式で近似することで、式(6)の非線形最適化問題を線形最適化問題の列に変換

する。すなわち、

$$\begin{aligned} \alpha^{(n+1)} &= \arg \min \{ E^{(n)}(\alpha) \} \quad (7) \\ E^{(n)}(\alpha) &= E_\alpha + \kappa'_c \sum_{(x,y)} W^{(n)}_F(x, y) (1 - \alpha(x, y))^2 \\ &\quad + \kappa'_c \sum_{(x,y)} W^{(n)}_B(x, y) (\alpha(x, y))^2 \quad (8) \end{aligned}$$

ここで  $n$  は反復の回数を示す。 $W^{(n)}_F(x, y)$ ,  $W^{(n)}_B(x, y)$  は  $(x, y)$  における  $n$  反復目の前景の重みと背景の重みであり、 $W^{(n)}_F(x, y)$  が大きければ  $\alpha(x, y)$  は 1 に偏り、 $W^{(n)}_B(x, y)$  が大きければ  $\alpha(x, y)$  は 0 に偏る。重みは式(6)の第二、第三項を現在の  $\alpha$  の見積もり  $\alpha^{(n)}$  で評価することで決定する。

$$\begin{aligned} W^{(n)}_F(x, y) &= \alpha^{(n)}(x, y) + \gamma (M_{B(\alpha^{(n)})}(x, y) - M_{F(\alpha^{(n)})}(x, y)) \\ W^{(n)}_B(x, y) &= 1 - \alpha^{(n)}(x, y) + \gamma (M_{F(\alpha^{(n)})}(x, y) - M_{B(\alpha^{(n)})}(x, y)) \quad (9) \end{aligned}$$

これは前景色ずれ度が背景色ずれ度より大きいときに前景の重みを下げ、背景の重みを上げることを意味する。逆も同様。  $\gamma$  は重みの上げ下げの強さを調整する係数である。初回は重みを 0 とする。

図 8 に反復に伴いマットが改善されていく様子を示す。図 8(a) に示す  $n = 0$  の結果は  $E_\alpha$  のみを最小化したことに相当するので、前景と背景の色が似ているところで誤抽出（赤矢印で示した）や十分に抽出しきれていない領域（緑矢印で示した）がある。反復により図 8(b)(c)(d) に示すようにこれらが修正されてゆく。

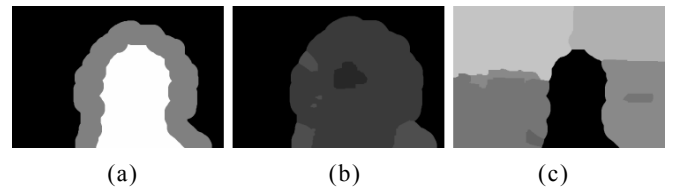


図 7: (a) 図 6(b) から生成したトライマップで、白が「確実に前景」、黒が「確実に背景」、灰色が「不明」領域を表す。(b) 補間した前景色ずれ量  $d_F(x, y)$ 。(c) 補間した背景色ずれ量  $d_B(x, y)$ 。

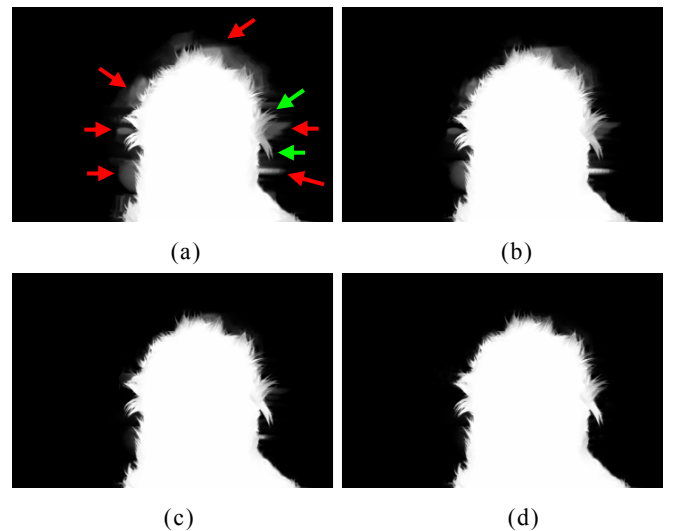


図 8: マット抽出計算の途中結果。(a) 初回の結果。(b) 1 回反復後。(c) 10 回反復後。(d) 20 回反復して収束した。

## 5. 結果

図 9 に屋外での人物写真への適用例を示す。装置が簡素なので屋外へ持ち出すことは容易である。図 9(b) に奥行き推定結果を示す。人物が手前にいることだけでなく、人物の左にある木の幹の形まで抽出できている。図 9(c)には抽出したマットを示す。髪や服の色が背景色に似ている難しい例題だが、比較的良好な抽出結果を得た。図 9(d)には異なる背景への合成例を示す。

図 10 に動物の写真への適用例を示す。提案法は一度の撮影で済むので、動き回る動物にも適用しやすい。色ずれ量とマットを利用した画像編集例をいくつか示す。図 11 二段目には背景の色ずれを相殺した後に前景と再度合成することで色ずれのない画像を復元した。背景の焦点ボケ除去を行ってから合成すると図 11 三段目の結果を得る。図 3 に示したようにボケ関数は菱形をしており、そのサイズは色ずれ量に一致するので、既知のボケ関数で逆畳み込みを行える。我々は[3]の手法を用いた。また、奥行きを利用して図 11 最下段のように少し視点をずらした画像の生成が可能である。

計算時間は  $720 \times 480$  の画像に対し Pentium 4 3.2GHz で奥行き推定が 30 秒、マット抽出が 10 分であった。使用したパラメータは  $\kappa_s = 0.001$ ,  $\kappa'_c = 0.01$ ,  $\gamma = 100.0$  である。局所ウィンドウ  $w(x, y)$  のサイズは  $15 \times 15$ , 取り得る視差の範囲は  $[-5, 10]$  とした (単位は画素)。

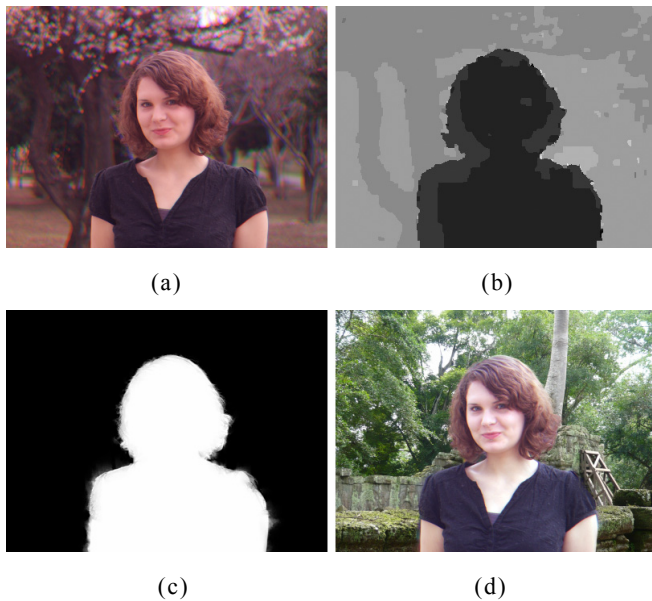


図 9: (a)撮影した写真。(b)奥行き推定結果。(c)前景マット抽出結果。(d)異なる背景への合成例。

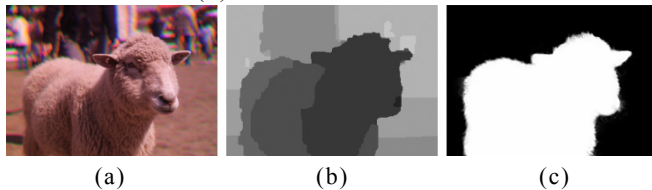


図 10: (a)撮影した写真。(b)奥行き推定結果。(c)前景マット抽出結果。



図 11: 上から順に、撮影画像 (図 4 と同じ)、色ずれを相殺した結果、背景の焦点ボケを除去した結果、視点を変更した結果。右には一部の拡大図を示した。

## 6. まとめ

カメラレンズの絞り位置に色フィルタを用いてシーンの奥行き推定と前景マット抽出を行う手法を提案した。既存のカメラに備わっている RGB 別の撮像機能を利用するためカメラ本体には加工が不要である。多くのカメラは一画素当たり RGB のうち一色を記録して残りを補間するので、解像度を犠牲にした多視点撮影[2,8,18]とみなすこともできるが、装置が簡素で露光も一度で済むことが提案法の利点である。提案法による色ずれのせいで色補間の性能が落ちる恐れがあるが、我々の実験した範囲では問題は認められなかった。

制限事項として、フィルタにより受光量が落ちることが挙げられる。開口部の分割数が 3 と少ないので、受光量のさらなる低下を避けるには RGB 各フィルタの面積を大きくする必要があり、画像には焦点ボケが含まれる。そのため、焦点の合った奥行きから離れるほどマッチングの精度が下がる。ただし、背景の高周波情報が低減されるのでマット抽出には有利にはたらく。また、マット抽出では前景と背景に奥行きのあることを仮定したので、地面に立っている人のように奥行きが連続的に変わる場合を扱えない。

今後の課題は画像間マッチング指標とマツト抽出性能の定量評価と、動画への適用である。また、マツト抽出の計算時間が長いので高速化を検討したい。

## 謝辞

貴重な助言をくださった苗村健准教授（東京大学）と井口雄介博士(NEC)に感謝します。

## 文献

- [1] Y. Amari, E. H. Adelson. "Single-eye range estimation by using displaced apertures with color filters," Proc. Int. Conf. Industrial Electronics, Control, Instrumentation, and Automation, vol.3, 1588-1592, 1992.
- [2] E. H. Adelson, J. Y. A. Wang. "Single lens stereo with a plenoptic camera," Trans. PAMI, Vol.14, No.2, pp.99-106, 1992.
- [3] Y. Bando, T. Nishita. "Towards digital refocusing from a single photograph," Proc. Pacific Graphics, pp.363-372, 2007.
- [4] Y. Boykov, O. Veksler, R. Zabih. "Fast approximate energy minimization via graph cuts," Trans. PAMI, Vol.23, No.11, 1222-1239, 2001.
- [5] I-C. Chang, C.-L. Huang, W.-J. Hsueh, H.-C. Lin, C.-C. Chen, Y.-H. Yeh. "A novel 3-D hand-held camera based on tri-aperture lens," Proc. SPIE 4925, 655-662, 2002.
- [6] Y.-Y. Chuang, B. Curless, D. H. Salesin, R. Szeliski. "A Bayesian approach to digital matting," Proc. CVPR, 264-271, 2001.
- [7] F. C. Crow. "Summed-area tables for texture mapping," Computer Graphics (Proc. SIGGRAPH), Vol.18, No.3, pp. 207-212, 1984.
- [8] T. Georgeiv, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, C. Intwala. "Spatio-angular resolution tradeoff in integral photography," Proc. Eurographics Symposium on Rendering, pp.263-272, 2006.
- [9] L. Goffman-Vinopal, M. Porat. "Color image compression using inter-color correlation," Proc. ICIP, Vol.2, pp.353-356, 2002.
- [10] P. Green, W. Sun, W. Matusik, F. Durand. "Multi-aperture photography," Trans. Graphics, Vol.26, No.3, Article 68, 2007.
- [11] N. Joshi, W. Matusik, S. Avidan. "Natural video matting using camera arrays," Trans. Graphics, Vol.25, No.3, pp.779-786, 2006.
- [12] A. Levin, D. Lischinski, Y. Weiss. "A closed form solution to natural image matting," Proc. CVPR, 61-68, 2006.
- [13] A. Levin, R. Fergus, F. Durand, W. T. Freeman. "Image and depth from a conventional camera with a coded aperture," Trans. Graphics, Vol.26, No.3, Article 70, 2007.
- [14] J. P. Lewis. "Fast template matching," Proc. Vision Interface, pp.120-123, 1995.
- [15] C.-K. Liang, G. Liu, H. H. Chen. "Light field acquisition using programmable aperture camera," Proc. ICIP, Vol.5, pp.233-236, 2007.
- [16] M. McGuire, W. Matusik, H. Pfister, J. F. Hughes, F. Durand. "Defocus video matting," Trans. Graphics, Vol.24, No.3, pp.567-576, 2005.

- [17] F. Moreno-Noguer, P. N. Belhumeur, S. K. Nayar. "Active refocusing of images and videos," Trans. Graphics, Vol.26, No.3, Article 67, 2007.
- [18] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, P. Hanrahan. "Light Field Photography with a Hand-Held Plenoptic Camera," Tech. Rep. CSTR 2005-02, Stanford Computer Science, Apr. 2005.
- [19] R. Ng. "Fourier slice photography," Trans. Graphics, Vol.24, No.3, pp.735-744, 2005.
- [20] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, J. Tumblin. "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing," Trans. Graphics, Vol.26, No.3, Article 69, 2007.
- [21] J. Wang, M. F. Cohen. "Optimized color sampling for robust matting," Proc. CVPR, pp.1-8, 2007.

## A. 付録

### A.1. 色変換

白い紙を RGB フィルタそれぞれ単独で観測したときの画像の色を  $k_r, k_g, k_b$  とする。3 節に述べた我々のプロトタイプでは以下の値であった。

$$k_r = (1.0, 0.335, 0.025)$$

$$k_g = (0.153, 1.0, 0.162)$$

$$k_b = (0.007, 0.190, 1.0)$$

ただし最大成分を 1 に正規化してある。よって  $M = (k_r, k_g, k_b)$  とすると、観測色  $c$  に対して  $M^{-1}c$  を計算すれば各フィルタを通った画像に分離できる。

### A.2. $L(x, y; d)$ と Normalized Cross Correlation

RG の二次元空間で式(1)に相当するものを考えると

$$L'(x, y; d) = \lambda_0 \lambda_1 / \sigma_r^2 \sigma_g^2$$

となる。R と G の共分散を  $\sigma_{rg}$  と書くと、 $\lambda_0 \lambda_1 = \det(\Sigma) = \sigma_r^2 \sigma_g^2 - \sigma_{rg}^2$  であるから、

$$L'(x, y; d) = 1 - \sigma_{rg}^2 / \sigma_r^2 \sigma_g^2$$

となる。右辺第二項は Normalized Cross Correlation であり、 $L'$  とは  $[0, 1]$  の大小関係が反転して対応している。

### A.3. $e(x, y; d)$ の導出

画素色の集合を  $\{c_i\}$ 、その平均を  $\mu$  とし、当てはめた直線を表す単位ベクトル（最大固有ベクトル）を  $l_0$  とすると、点  $c_i$  の直線  $l_0$  からの距離  $t_i$  は

$$t_i^2 = |c_i - \mu|^2 - ((c_i - \mu)^T l_0)^2$$

であるから、距離  $t_i$  の二乗平均は

$$\begin{aligned} e(x, y; d) &= \sum_i |c_i - \mu|^2 / n - \sum_i ((c_i - \mu)^T l_0)^2 / n \\ &= \sigma_r^2 + \sigma_g^2 + \sigma_b^2 - l_0^T (\sum_i (c_i - \mu) (c_i - \mu)^T / n) l_0 \\ &= \sigma_r^2 + \sigma_g^2 + \sigma_b^2 - l_0^T \Sigma l_0 \end{aligned}$$

ここで  $l_0$  は  $\Sigma$  の固有値であるから  $l_0^T \Sigma l_0 = \lambda_0 l_0^T l_0 = \lambda_0$ .