

多視点画像の相互色較正

木村優太[†] 延原章平[†] 松山隆司[†]

[†] 京都大学情報学研究科

〒 606-8501 京都市左京区吉田本町

E-mail: †{kimura,nob,tm}@vision.kuee.kyoto-u.ac.jp

あらまし 3次元形状復元など、撮影対象を複数の異なる視点から同時に撮影する状況下では、カメラの撮像素子やレンズ特性の個体差などの理由で、例えば撮影された3次元点が完全拡散面であったとしても観測される色が異なる。また特にカメラが疎で光軸の方向が互いに大きく異なるとき、撮影画像間に単純な重なりが存在することは期待できず、従来の色校正を適用することは困難である。そこで本研究では同じ対象を異なる視点から撮影しているという状況を利用して、対象自身を用いて色ずれを校正する手法を提案する。本手法によって例えばインターネット上から収集した画像を用いた3次元形状復元のように、本来の撮影環境が未知の状況においても色校正を行うことが可能となる。
キーワード 多視点画像, カメラ色較正, テクスチャマッピング

Color Calibration for Multi-viewpoint Cameras in Sparse and Convergent Arrangement

Yuta KIMURA[†], Shohei NOBUHARA[†], and Takashi MATSUYAMA[†]

[†] Graduate School of Informatics, Kyoto University, Japan

Yoshidahonmachi, Sakyo, Kyoto, Japan

E-mail: †{kimura,nob,tm}@vision.kuee.kyoto-u.ac.jp

Abstract This paper is aimed at proposing a new color calibration method for multi-viewpoint images captured by sparsely and convergently arranged cameras. While most of conventional approaches depend on the overlapping areas on each images, sparse and convergent camera arrangement does not provide such overlapping. We propose a new method which estimates the color calibration parameters using the 3D object in the scene. Our method estimates the 3D points in the scene by a robust wide-baseline stereo and then utilize their positions to define an error function which measures color differences between images. Finally a set of parameters which minimizes this error by a non-linear optimization method. Experiments with real images show that our method can minimize the difference of pixel values (1) quantitatively by leave-one-out evaluation, and (2) qualitatively by rendering a 3D video.

Key words multi-viewpoint images, color calibration, texture mapping

1. はじめに

多視点画像の色較正はコンピュータビジョンの分野では古くから取り組まれてきた問題の一つであり、特に近年では口径蝕と撮像素子特性を同時に校正する手法が提案されている [1] ~ [4]。しかしこれらはいずれもパノラマ画像の生成のように、各画像は (1) 互いに光軸が平行に近い状態で、かつ (2) 1つのカメラが移動して撮影される状況を想定している。このため (1) 画像間には単純に重なりが存在し、(2) レンズや撮像素子の個体差を考慮する必要がないことを仮定できた。

これに対して本論文では、図 1 に示すように疎でかつ

対象を囲むような、つまり光軸が対象に向けて収斂するようなカメラ配置を想定する。また運動する物体の3次元形状復元 [5], [6] や、インターネット上から収集された画像を用いた3次元形状復元 [7] を想定して各カメラは物理的に異なるとする。このため撮影画像は互いに単純な重なりを持たず、また各カメラの色較正パラメータはそれぞれ個別に推定するものとする。

このような問題に対して、我々は (1) まず多視点画像からテクスチャ情報に基づいて対象表面である可能性が高い3次元点を推定し、(2) 推定された点を観測可能なカメラにおける観測輝度値が互いに一致するような色較正パラメータを推定するというアプローチを採る。なお

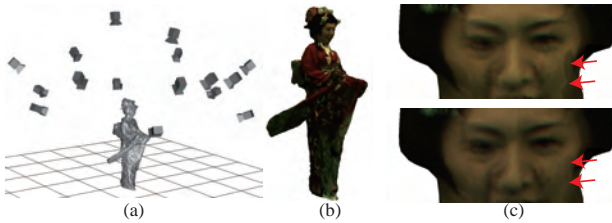


図 1 疎な多視点画像の色校正．(a) 対象を囲むカメラ配置，(b) 復元された形状にテクスチャを貼った様子，(c) 顔周辺の拡大図．上：テクスチャに用いたカメラ番号による色分け，中：色校正無，下：色校正有．赤矢印はテクスチャ境界を指す．

本論文では，3次元形状復元そのものを行うことができる一方で色校正が行えていないという状況を想定しているため，カメラの内部および外部パラメータは既知であるとする．また多視点画像間で対応付けを行って得られる点は完全拡散反射面であると仮定する．

以降では関連研究に対する位置づけの後に，多視点画像間の頑健な対応付け方法について述べ，次に観測過程のモデル化について説明する．そしてこのモデルパラメータを推定する方法を述べた後に，実験によって提案手法の有効性を示す．

2. 関連研究

前節において述べたように，従来の多視点画像から口径蝕と撮像素子特性を同時に推定する手法 [2] ~ [4] においてはカメラの光軸が互いに平行に近い配置を想定し，またその結果として画像間に単純な重なりが存在することを期待している．そのため図 1(a) のように光軸が対象に向けて収斂し，かつ疎な配置を想定している我々の問題に直接適用することは難しい．

またカメラ配置とは関係なく，単純に 1 枚の撮影画像から口径蝕を推定する手法も提案されている [8]．しかしこの手法では画像中心からの放射状の濃度変化の原因が口径蝕が支配的であることを仮定するため，室内の撮影スタジオ環境のように近接光源に照らされた背景が存在する場合には有効に機能しない．

3. 多視点画像における頑健な特徴点对应推定

本節で導入するアルゴリズムの目的は，後の色校正パラメータ推定に使用する誤差関数を定義する上で必要となる多視点画像間の対応点データを得ることである．基線長が長く，かつ光軸が平行ではない場合の頑健な特徴点抽出および対応付けにはこれまでも多くの手法が提案されてきたが [9] ~ [11]，ここではカメラ配置が既知であることを前提に [12] の手法に一意性の制約を加えた以下の手法を提案する．

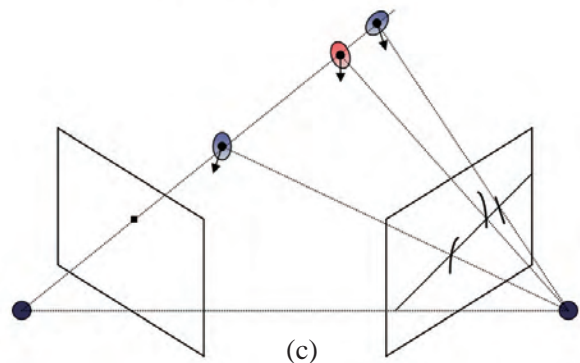
- (1) 2つのカメラ c と c' について撮影画像それぞれにおいてエッジ検出を行う．
- (2) このエッジ画像からエピポーラ線と交差するエッ



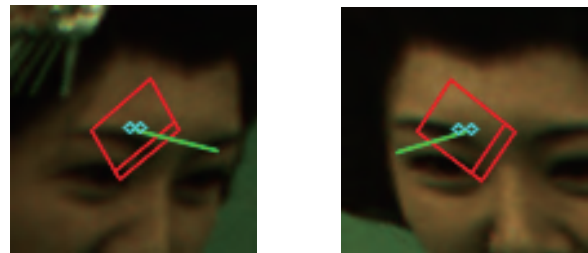
(a)



(b)



(c)



(d)

図 2 エッジ情報を用いた対応付け．(a) 平行化された画像対，(b) エピポーラ線と交差するエッジ成分，(c) 法線の最適化を伴うテクスチャ不一致度の計算．(d) 対応づけられた特徴点の例．(d) の赤枠はテクスチャ不一致度の計算に用いたウィンドウの大きさを，緑色の線は推定された法線方向を，水色の丸はエピポーラ線に沿ってエッジを挟む 2 点を示している．

ジ成分以外を除去する．こうして得られた画像を I_E, I'_E とする (図 2(b))．この段階で I_E に含まれるエッジ上の点 e に対して， C' のエピポーラ上には複数のエッジが交差する．この交点の集合を $E' = \{e'_j | j = 1, \dots, n\}$ とする．

(3) e とそれに対応する $e'_j \in E'$ について, n 個の組み合わせでそれぞれテクスチャの相違度を計算する. このとき [12] と同様に法線方向の最適化も行う. こうして 1 つの e に対して n 個の相違度が求まる. ここで最も低い相違度を与えた E' の要素を e'_j とすると, e と e'_j の間の相違度が, 2 番目に低い相違度に対して十分小さいとき, e と e'_j をエッジ対応点として採用する (図 2(c)). ここで相違度 ρ の計算は口径蝕や撮像素子特性による影響に対して頑健である必要がある. 本論文ではこれを満たすものの 1 つとして知られる Zero-mean Normalized Cross Correlation 関数 ZNCC(\cdot) を用いて

$$\rho = 1.0 - \text{ZNCC}(e, e'_j) \quad (1)$$

として相違度を計算する.

(4) こうして得られたエッジ対応点は, その定義から画像のエッジ上に存在する. そのためその座標における輝度値がエッジを挟む明暗どちらとなるかは安定ではない. そこでエッジ線に沿ってこのエッジを挟む 2 点を, 後の色校正に使用する対応点として用いる (図 2(d) の水色の丸で囲まれた点). こうして得られた対応点の 1 つの組を $\langle p_c, p_{c'} \rangle$ と記述することとし, カメラ c と c' の間で得られた $n_{c,c'}$ 個の対応点の集合を $P_{c,c'} = \{\langle p_c^i, p_{c'}^i \rangle \mid i = 1, \dots, n_{c,c'}\}$ とする. 以上をすべてのカメラの組み合わせについて行い, それぞれで対応点集合を得る.

4. 撮影過程のモデル化

次に撮影過程をモデル化し, 本研究で用いた色校正関数の具体的な内容について述べる. シーン中の完全拡散反射面から放射された光 l が撮像素子で受光されるまでの過程は次の式 (2) で表されるとする.

$$l^i = G^i(V(l, r)), \quad i = \{R, G, B\} \quad (2)$$

ただし $G^i(l)$ は撮像素子の R, G, B 各チャンネルのゲイン特性を表す関数であり, $V(l, r)$ はレンズのシェーディング特性を表す関数である. また r は撮像面上において, 光学中心から光が通過する位置までの距離を表す. つまりある点から放射された光は口径蝕の影響を受けてカメラ画像の周辺部に向かうにつれ減衰し, さらに受光素子で各撮像素子の影響を受けて画素値に変換されるとモデル化する.

4.1 レンズ特性モデル

本論文では, 画像の光学中心から r だけ離れたカメラ画像上の点の口径蝕関数 $V(l, r)$ が以下の式 (3) で表される簡略化された Kang-and-Weiss モデル [13] を用いる.

$$V(l, r) = V_G(r)V_A(r)l \quad (3)$$

$$V_G(r) = (1 - ar) \quad (4)$$

$$V_A(r) = \frac{1}{(1 + (r/f)^2)^2} \quad (5)$$

ただし $V_G(r)$ は口径蝕を, $V_A(r)$ はコサイン 4 乗則を表し, f はカメラの有効焦点距離, a は $V_G(r)$ の 1 次の係数である. 本論文ではカメラの光学中心は内部パラメータから既知であると仮定するが, 内部パラメータの焦点距離と式 (3) の有効焦点距離 f は別のものとして扱う. したがってレンズ特性モデルの未知パラメータはカメラ 1 台に対して a_0 と f の 2 個である.

4.2 撮像素子特性モデル

本論文では R, G, B 各チャンネルそれぞれが, 独立な線形ゲインモデル

$$G^i(l) = \alpha^i l + \beta^i, \quad i = \{R, G, B\} \quad (6)$$

に沿って画素値を出力するとモデル化する. したがって撮像素子特性モデルの未知パラメータはカメラ 1 台に対して 6 個である. ただし以降では特に必要のない限りチャンネルを表す添え字 $\{R, G, B\}$ は省略する.

5. 提案手法

ここまでで, (1) 口径蝕補正モデルと (2) 撮像素子特性モデルを定義した. 本節ではまずこのモデルから観測輝度値を理想的な輝度値へと変換する色校正関数を定義する. そして第 3. 節で求めた対応点間で, 理想輝度値が一致するという仮定のもとに誤差関数を定義し, この誤差関数の値が最小化するようなモデルパラメータを非線形最小二乗問題として解くことで推定する.

5.1 色校正関数

第 4. 節で述べたモデルを l について逆に解いて, 口径蝕と撮像素子特性によって生じる撮影画像の輝度変化を同時に補正する関数 F を導出する. あるカメラ c で撮影された画像上のある点で計測される画素値を x とすると, カメラ c の色校正関数 $F(x)$ は次式になる.

$$F_c(x) = \frac{(x - \beta_c)(1 + (r/f_c)^2)^2}{\alpha_c(1 - a_c r)} \quad (7)$$

ここで f_c はカメラ c の有効焦点距離, a_c は口径蝕パラメータ, α_c, β_c はカメラ c のゲイン特性に関するパラメータであり, $\{R, G, B\}$ 各チャンネル独立に存在するとモデル化する.

5.2 誤差関数

第 3. 節で求めたカメラ c と c' 間の対応点集合 $P_{c,c'}$ に含まれる各対応点 $\langle p_c^i, p_{c'}^i \rangle$ について, 先に述べた色校正関数 $F(x)$ を用いて理想輝度値を計算すると, その差 $E(\langle p_c^i, p_{c'}^i \rangle)$ は

$$\begin{aligned}
E(\langle p_c^i, p_{c'}^i \rangle) &= \rho \left(F_c(x_{p_c^i}) - F_{c'}(x_{p_{c'}^i}) \right) \\
&= \rho \left(\frac{(x_{p_c^i} - \beta_c)(1 + (r/f_c)^2)^2}{\alpha_c(1 - a_c r)} \right. \\
&\quad \left. - \frac{(x_{p_{c'}^i} - \beta_{c'}) (1 + (r/f_{c'})^2)^2}{\alpha_{c'}(1 - a_{c'} r)} \right) \quad (8)
\end{aligned}$$

となる．ただし $x_{p_c^i}$ および $x_{p_{c'}^i}$ はそれぞれ p_c^i および $p_{c'}^i$ における観測輝度値であり， $\rho(\cdot)$ はロバスト推定用の関数である．本論文では

$$\rho(\delta) = 2b^2 \left(\sqrt{1 + \frac{\delta^2}{b^2}} - 1 \right) \quad (9)$$

で定義される pseudo-Huber 関数 [14] を使用し， b の値は人手によって決定した．

これを用いるとあるカメラ c と c' 間の $n_{c,c'}$ 個の対応点集合 $P_{c,c'}$ から， $n_{c,c'}$ 次元の誤差ベクトル

$$E_{c,c'} = \left(E(\langle p_c^1, p_{c'}^1 \rangle), \dots, E(\langle p_c^{n_{c,c'}}, p_{c'}^{n_{c,c'}} \rangle) \right)^\top \quad (10)$$

が定義できる．これを N 台のカメラすべての組み合わせ ${}_n C_2$ 通りについて定義し，すべてを結合することで全対応点に関する誤差ベクトルを

$$E_0 = (E_{1,2}^\top \quad E_{1,3}^\top \quad \dots \quad E_{N-1,N}^\top)^\top \quad (11)$$

として定義する．

しかしこの E_0 を最小化するモデルパラメータには， $\alpha_c = \infty$ という自明な解が存在する．そのためこのままでは意味のあるパラメータを得ることはできないため，以下のように元の観測誤差に対する忠実度を導入する．

$$D_c(p_c^j) = \left\| x_{p_c^j} - F_c(x_{p_c^j}) \right\|^2 \quad (12)$$

ただしここで p_c^j はカメラ対に関わりなくカメラ c で得られた全ての対応点であるとする．カメラ c で得られた対応点の総数を n_c とすると， c における忠実度ベクトルは

$$D_c = (D_c(p_c^1), \dots, D_c(p_c^{n_c}))^\top \quad (13)$$

となり，カメラ全体として

$$D_0 = (D_1^\top \quad D_2^\top \quad \dots \quad D_N^\top)^\top \quad (14)$$

と定義できる．以上の E_0 および D_0 を同時に最小化するために，最終的な誤差ベクトルを

$$E = (E_0^\top \quad \lambda D_0^\top)^\top \quad (15)$$

とする．ただし λ は元の観測輝度値への忠実度に関する重み係数である．

以上から， N 台のカメラが持つ $8N$ 個の未知パラメータに対して，誤差値 E が式 (15) のように定義された．本論文では各カメラにおいて 8 個以上の対応点が得られると仮定して， E を最小化するパラメータを修正 Levenberg-Marquardt 法で求める．ただし各パラメータの初期値は $\alpha_c = 1$ ， $\beta_c = 0$ ， $a_c = 1$ とし， f_c に関しては内部パラメータの焦点距離の値を用いた．

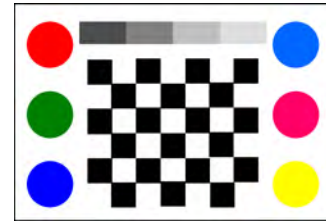


図 3 既知の参照物体

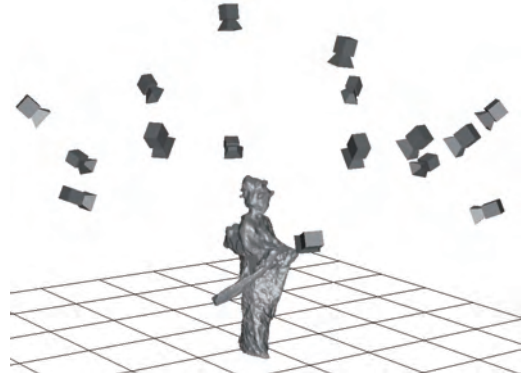


図 4 カメラ配置

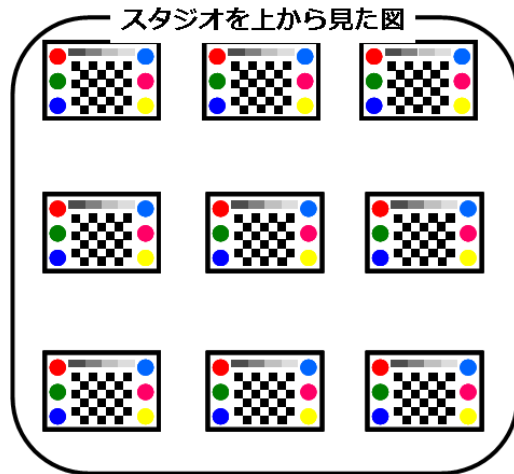


図 5 参照物体の配置

6. 評価実験

6.1 参照物体を用いた評価

まず多視点画像間の対応付けを除いた，パラメータ推定のみ妥当性を評価するために，図 3 のような既知の物体を用いた実験を行った．カメラは図 4 に示すように対象を囲むように配置された 16 台の UXGA カメラ (Pointgrey 製 Grasshopper, レンズ Kowa LM5JC1M) を使用した．この実験では図 3 の物体を図 5 のように撮影空間中を均等にカバーするように 9 回に分けて撮影を行って対応点を取得した．撮影された画像の一部を図 6 に示す．

この入力画像からパラメータ推定を行い，画像の補正を行った結果が図 7 である．同図左上が補正前の入力画像の 1 つであり，右上がそれを補正した画像である．ま

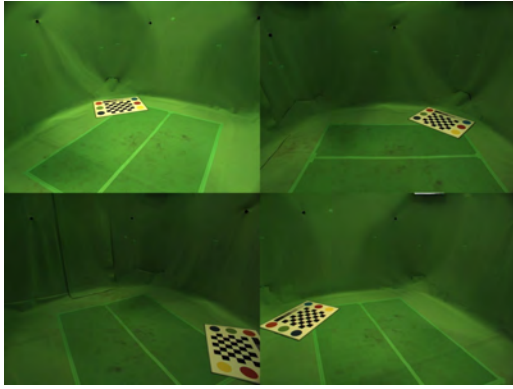
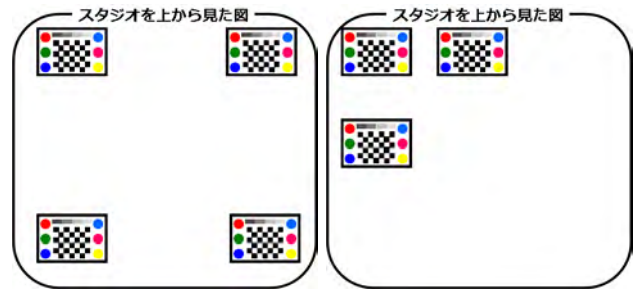


図 6 参照物体を撮影した多視点画像の一部



(a)

(b)

図 9 対応点分布に偏りのある場合

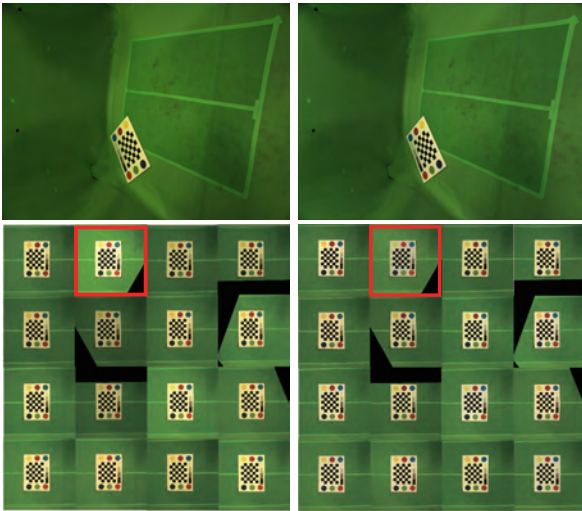


図 7 色補正結果．左：補正前，右：補正後，上：カメラ画像の例，下：ホモグラフィによって同一仮想視点へと変換された全カメラの画像．

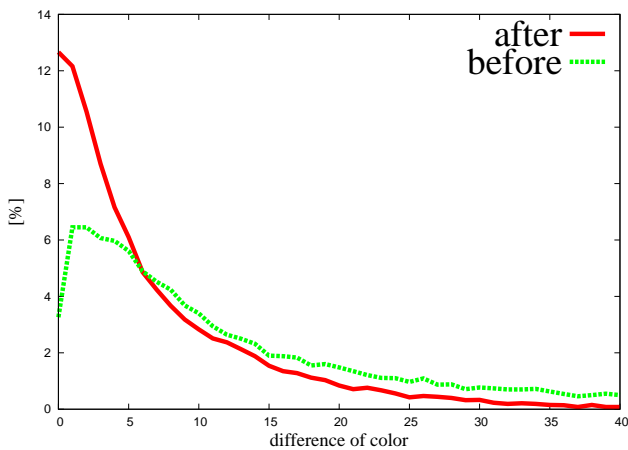
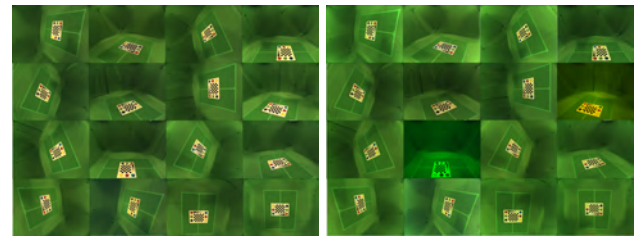


図 8 色補正前後における対応点間の画素値の相違度

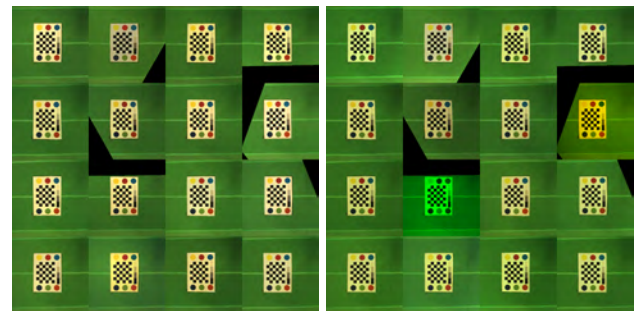
た同図左下および右下はホモグラフィによって全 16 台のカメラ画像を同一仮想視点へと変換した結果である．この結果を比較すると，特に赤く囲んだ視点において，補正によって色の一致度が向上していることが定性的に確認できる．また図 8 に各対応点における補正前後の輝度差分布のグラフを示す．横軸は輝度差，縦軸は対応点



(a) 図 9(a) の場合

(b) 図 9(b) の場合

図 10 対応点分布に偏りのある場合の補正結果



(a) 図 9(a) の場合

(b) 図 9(b) の場合

図 11 対応点分布に偏りのある場合の補正結果（仮想視点に投影した結果）

表 1 参照点の空間分布による対応点間の画素値の差の平均値への影響

較正前画像	図 5	図 9(a)	図 9(b)
	15.60	8.532	11.42

全体に占める割合であり，補正によって全体として輝度差が小さくなっていることが定量的に確認できる．

次に対応点の分布による影響に関する実験として，図 9 のような配置におけるパラメータ推定を行った．同図 (a) は空間に均等に分布する一方に対応点の数は少なくなっており，同図 (b) はさらに明確に空間的に偏りが存在する．このような 2 つの場合の補正結果を図 10 およ

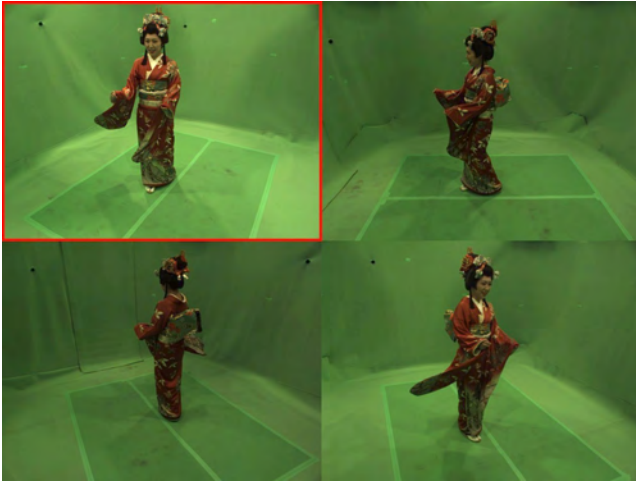


図 12 対象を撮影した多視点画像



図 13 推定された対応点から得られた 3次元点の分布

び図 11 に、また補正前後における対応点間の輝度差の平均値を表 1 に示す．図 10, 11 から (a) 対応点の密度の低下の影響はあまり見られない一方で、(b) 対応点の分布に偏りがある場合は明らかに不適当なパラメータが推定されてしまい、補正後の画像の色が明らかにおかしくなっていることが定性的に確認できる．また対応点間の画素値の残差についても、表 1 に示すように分布に偏りがあることによって悪化することが定量的に確認できる．

6.2 対象自身から得られる対応点を用いた評価

次に対象自身から得られる対応点を用いた評価を行う．まず撮影画像を図 12 に示す．また第 3. 節のアルゴリズムで得られた対応点集合から三角測量によって計算した 3次元点群を図 13 に示す．この実験では 16 台のカメラで合計 67,382 個の対応点を得られた．

この対応点を用いて色校正パラメータを推定し、得られたパラメータで画像を補正した結果を図 14 に示す．また比較のため、前節で既知の参照物体を使用して求めた色校正パラメータで画像を補正した結果を図 15 に示す．

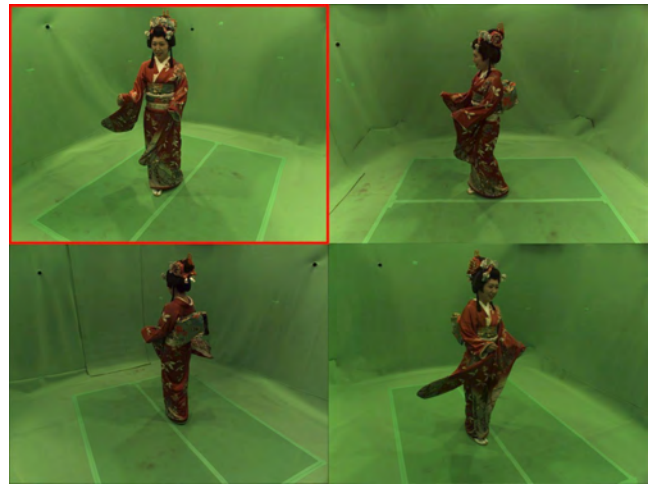


図 14 対象上の対応点から推定されたパラメータによる補正結果

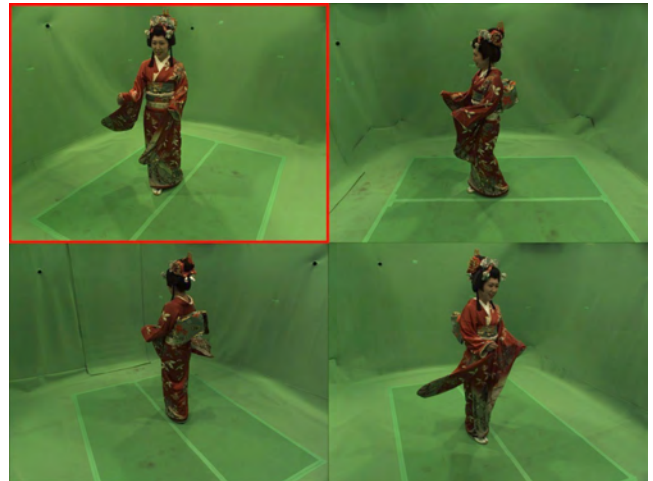


図 15 既知の参照物体から推定されたパラメータによる補正結果

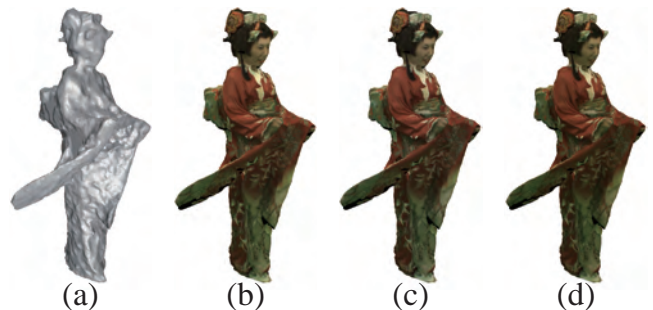


図 16 補正後の多視点画像を用いたテクスチャマッピング結果．(a) 推定された 3次元形状 [6]、(b) 元画像を用いた場合、(c) 既知の参照物体から推定されたパラメータで補正した画像を用いた場合、(d) 対象上の対応点から推定されたパラメータで補正した画像を用いた場合．

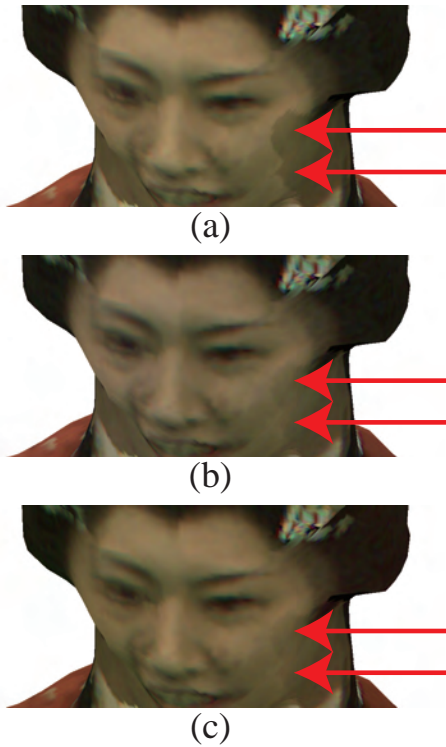


図 17 補正後の多視点画像を用いたテクスチャマッピング結果(拡大図)(a)元画像を用いた場合,(b)既知の参照物体から推定されたパラメータで補正した画像を用いた場合,(c)対象上の対応点から推定されたパラメータで補正した画像を用いた場合. 矢印はそれぞれテクスチャマッピングに使用するカメラが切り替わる境界を指す.

特に赤枠で囲まれた視点について図 12, 15, 14 を比較すると, 撮影対象から対応点を求めた場合においても, 既知の参照物体を使用した場合と同様に色校正が行えていることが定性的に確認できる. またこうして補正された画像を用いて [6] の手法で復元された 3 次元形状に対してテクスチャマッピングを施した結果を図 16 および図 17 に示す. 図 17 の矢印はテクスチャとして採用するカメラが切り替わる境界を指しており, ここに注目すると元画像を用いた場合(図 17(a))に比べて, 撮影対象から対応点を求めた場合(図 17(c))が既知の参照物体を使用した場合(図 17(b))と同様に色校正によってテクスチャ境界がより滑らかになっていることが確認できる. 以上から, 撮影対象から対応点を自動的に求めた場合においても, 既知の参照物体を使用した場合と同様に色校正を行うことができることが確認できた.

7. 結 論

本論文では, 対象を囲むように配置された疎な多視点カメラ群の色校正手法として, 撮影対象自身を色校正基準物体として使用する手法を提案した. 提案手法のアプローチは決められた基準物体を事前に撮影する必要があるなどの制約が無いため, インターネット上から収集された多視点画像のように, 撮影時の環境が未知の場合に

においても適用が可能であるという利点がある

しかし一方で, 本論文では (1) 得られる対応点は全て完全拡散面であると仮定しており, また (2) ガンマ関数に関する推定を行っていない. 今後はこのような点について制約の緩和を行い, より適用可能性の高い手法を検討する予定である.

文 献

- [1] M. D. Grossberg and S. K. Nayar: "Modeling the space of camera response functions", *PAMI*, **26**, 10, pp. 1272–1282 (2004).
- [2] D. B. Goldman and J.-H. Chen: "Vignette and exposure calibration and compensation", *Proc. of ICCV*, pp. 899–906 (2005).
- [3] A. Litvinov and Y. Y. Schechner: "Addressing radiometric nonidealities: A unified framework", *Proc. of CVPR*, pp. 52–59 (2005).
- [4] S. J. Kim and M. Pollefeys: "Robust radiometric calibration and vignetting correction", *PAMI*, **30**, 4, pp. 562–576 (2008).
- [5] J. Starck, A. Hilton and G. Miller: "Volumetric stereo with silhouette and feature constraints", *Proc. of BMVC* (2006).
- [6] T. Tung, S. Nobuhara and T. Matsuyama: "Simultaneous super-resolution and 3d video using graphcuts", *Proc. of CVPR*, pp. 1–8 (2008).
- [7] N. Snavely, S. M. Seitz and R. Szeliski: "Photo tourism: exploring photo collections in 3d", *Proc. of SIGGRAPH*, pp. 835–846 (2006).
- [8] Y. Zheng, J. Yu, S. Kang, S. Lin and C. Kambhamettu: "Single-image vignetting correction using radial gradient symmetry", pp. 1–8 (2008).
- [9] D. G. Lowe: "Distinctive image features from scale-invariant keypoints", *IJCV*, **60**, 2, pp. 91–110 (2004).
- [10] K. Mikolajczyk and C. Schmid: "Scale & affine invariant interest point detectors", *IJCV*, **60**, 1, pp. 63–86 (2004).
- [11] E. Tola, V. Lepetit and P. Fua: "A fast local descriptor for dense matching", *Proc. of CVPR*, pp. 1–8 (2008).
- [12] Y. Furukawa and J. Ponce: "Accurate, dense, and robust multi-view stereopsis", *Proc. of CVPR*, pp. 1–8 (2007).
- [13] S. B. Kang and R. S. Weiss: "Can we calibrate a camera using an image of a flat, textureless lambertian surface?", *Proc. of ECCV*, pp. 640–653 (2000).
- [14] R. I. Hartley and A. Zisserman: "Multiple View Geometry in Computer Vision", Cambridge University Press (2000).