

光線方向の違いを考慮したライトフィールド特徴量の提案*

—物体形状とカメラ姿勢に頑健な画像対応付け—

清水政行** 川西康友** 出口大輔** 井手一郎** 村瀬 洋**

Proposal of a Light-Field Descriptor Considering Light-Ray Direction
—Image Matching Robust to Object Shape and Camera Pose—

Masayuki SHIMIZU, Yasutomo KAWANISHI, Daisuke DEGUCHI, Ichiro IDE and Hiroshi MURASE

Many camera and photo applications, for example image mosaicing, Simultaneous Localization And Mapping (SLAM), and so on, use image keypoints, descriptors, and their matching. Recently, obtaining light-field has become easier thanks to the commercialization of light-field cameras. Light-field is a record of not only light ray intensity but also directions, and multi-view images can be acquired with it. In this paper, a light-field descriptor with light ray directions using light-field is proposed. A light-field descriptor is defined as a set of feature vectors that change according to the direction of the light ray. In addition, a method to match these descriptors is proposed. As the method focuses on the direction of the light ray, it allows more accurate matching than a conventional method when the camera pose changes against the target. To confirm the effectiveness, the proposed method is evaluated by using light-field image recorded in a real scene. As a result, the proposed method is shown to improve the accuracy over a conventional method. In particular, the closer the distance between the target and the camera, the better accuracy of the proposed method could be achieved.

Key words: light-field, keypoints, descriptor, matching, focal length, ray direction

1. 緒 言

カメラ画像を対象とした認識技術は日々発展しており、その中でも特徴点、局所特徴量の記述及びその対応付けは、イメージモザイク¹⁾、同一画像検索²⁾など幅広く活用されている。近年では、コンピュータ性能の向上や計算コスト削減などにより、Simultaneous Localization And Mapping (SLAM)³⁾など実時間性が要求されるような応用でも活用されるようになってきた。

局所特徴量の記述に関するこれまでの研究を分類すると、画像変換に対する不変性の向上⁴⁾、特徴量の識別力の向上⁵⁾、特徴量の2値化等による計算の高速化と省メモリ化^{6) 7)}に分けられる。特徴点の対応付けにおいて、画像変換に対する不変性は重要であり、スケール変化、回転やAffine変換に不変な特徴点の抽出と特徴量の記述手法が提案されている^{4) 8) 9)}。これらの手法は被写体が平面物体であると仮定しているが、立体的な被写体に対してカメラ視点が変化した画像間で特徴点を対応付けする場合、この仮定は成立せず、特徴点の対応付け性能が低下するという課題がある。この課題を解決して対応付け性能を向上させるためには、視点変化に対して頑健な特徴量が必要になる。

近年、光線情報(ライトフィールド)を画像(以降、ライトフィールド画像と呼ぶ)として記録できるカメラが注目されている。初期の研究では、多数のカメラを縦横に並べたカメラレイによってライトフィールド画像を記録するライトフィールドカメラが用いられていた¹⁰⁾。最近では、撮像素子の前面にマ

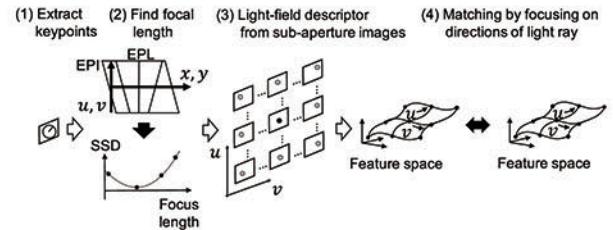


Fig. 1 Overview of light-field descriptor matching

イクロレンズアレイを配置することで密なカメラアレイを実現したライトフィールドカメラが製品化されており^{11) 12)}、ライトフィールド画像の記録が容易になった。ライトフィールド画像からは様々な方向の光線情報を得ることが可能であり、部分開口画像として様々な視点の画像を生成することができる¹³⁾。本論文ではこの性質をふまえ、視点変化に伴う特徴量の変化を表現可能な局所特徴量として「ライトフィールド特徴量」を定義し、特徴点の対応付け手法を提案する。具体的には、部分開口画像から局所特徴量を抽出し、光線方向の変化に応じて変化する特徴ベクトルの集合としてライトフィールド特徴量を定義する。ライトフィールド特徴量の対応付けは、同じ光線方向に対応するベクトル間の L_2 距離に基づいて行なう。これにより、同一の点に対して、同じ方向から撮影した部分開口画像で算出した特徴量で対応付けすることが可能になり、視点変化に伴う画像間の特徴点の対応付けの精度が向上する。ライトフィールド特徴量は、ライトフィールドから実際に視線変化させた時の特徴量を算出できるため、従来手法に比べて、被写体が立体物であっても対応付けの精度が低下しないという利点がある。

図1に提案手法の概要を示す。本手法は、(1)特徴点の抽出、(2)抽出された特徴点毎に光線方向の変化による視差の推定、

* 原稿受付 令和2年5月22日

掲載決定 令和2年10月16日

** 名古屋大学(愛知県名古屋市中千種区不老町)

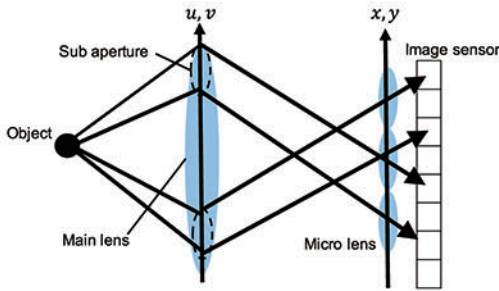


Fig. 2 Principle of light-field information recording

(3) 抽出された特徴点と推定された視差を用いたライトフィールド特徴量の抽出, (4) 光線方向を考慮した対応付け, から構成される。

本論文の貢献は以下の通りである。

1. ライトフィールド特徴量を用いた特徴点对応付け手法の提案
2. ライトフィールド特徴量が高い効果を発揮する撮影条件の明確化

以降, 2章で関連研究について述べる。3章で広く用いられている SIFT について説明した後, 4章でライトフィールド特徴量とその対応付け方法を説明する。そして, 5章で評価実験の詳細を述べるとともに, 6章で考察を加える。最後に7章で本論文をまとめる。

2. 関連研究

特徴点の対応付けに関連する研究は数多くなされているが, なかでも, Lowe らが提案した Scale-Invariant Feature Transform (SIFT) ⁴⁾ は最も有名な手法の1つである。SIFT はスケール, 回転, 照明変化に対して不変な特徴量である。さらに Speeded-Up Robust Feature (SURF) ¹⁴⁾ や PCA-SIFT ⁵⁾ など, SIFT を改良した手法が数多く提案されている。しかし, これらの手法はカメラの視点変化に対して不変でないという課題がある。この問題に対して, Mikolajczyk らは特徴点とその周辺画素値の2次の微分行列から視点変化を表現する行列を推定し, 特徴点を中心とした画像パッチを Affine 変換することで視点変化に対し頑健な特徴量を抽出する手法を提案している ⁹⁾。Morel らが提案している A-SIFT では入力画像から様々な Affine 変換した画像を複数生成して特徴点を抽出し, 総当たりでこれらの対応付けを行なっている ¹⁵⁾。しかし, これらの手法は撮影対象が平面と仮定して変換を行なっているため, 立体的な被写体に対するカメラの視点変化の不変性を解決できていない。

ライトフィールドカメラは空間中の光線情報を取得することが可能なカメラである。Ng らは, 撮像素子の前面にマイクロレンズアレイを配置することで密なカメラアレイを実現したライトフィールドカメラを提案しており ^{16) 17)}, Lytoto 社 ¹¹⁾ から製品化されている。図2は, ライトフィールドカメラの撮像の様子を図示したものである。物体が発する光は主レンズの部分開口とマイクロレンズを通り, 撮像素子の各画素に記録される。この機構によって, マイクロレンズ面での座標 (x, y) と主レンズ面の座標 (u, v) を通過する光の光量 $L(x, y, u, v)$ を記録することができる。これをライトフィールドと呼ぶ。ライトフィールドからは, 同一方向の光線を集めて部分開口画像を生

成することが可能であり, 式(1)のように表される。

$$I_{u,v}(x, y) = L(x, y, u, v) \quad (1)$$

ここで光線方向は, $u \in U, v \in V$ と x, y との組み合わせにより定まる。 U, V はライトフィールドカメラが記録できる光線方向の範囲であり, ライトフィールドカメラのレンズ設計により決まる値である。特に, $u = 0, v = 0$ の時, 主レンズ中央を投影中心とした透視投影カメラで撮像した画像を取得することができて, これを中心画像 $I_{0,0}$ と呼ぶ。ライトフィールドから得られる部分開口画像を活用した撮影後のリフォーカス ¹⁶⁾ や深度推定 ¹⁸⁾ 等の手法が多数提案されている。我々もこれまでにライトフィールドを活用することで, SIFT 特徴等による画像対応付けの精度向上を図る手法を提案している ¹⁹⁾。ライトフィールドから特徴点の焦点距離を推定し, 特徴点周りの局所画像に焦点がもっとも合った画像を使って特徴量を抽出することで画像対応付け精度が向上することを確認している。Lumentut らは各光線方向から生成された部分開口画像から特徴点を抽出し, 総当たりで対応付けを行なう手法を提案している ²⁰⁾。また, Tosic らはスケール空間だけでなく, EpiPolar Image (EPI) に対してもフィルタ処理を施し, 奥行きを考慮した特徴点の抽出と奥行きを推定を同時に行なう手法を提案している ²¹⁾。従来技術はライトフィールドを活用した特徴点の抽出に関する提案であるが, 本論文は, 光線情報を活用した, カメラの視点変化による特徴量変化の表現を可能にするライトフィールド特徴量とその対応付け手法について提案する。

3. SIFT を用いた特徴点对応付け

本論文で提案する手法は, 光線方向を考慮可能なライトフィールド特徴量へと SIFT 特徴量を拡張するものである。そのため, まずは本章で SIFT について説明する。SIFT は局所画像のスケール, 回転, 照明変化に対して頑健に対応付けすることが可能なアルゴリズムであり, その処理を (1) 特徴点の抽出, (2) 特徴量の算出, (3) 特徴点の対応付けの3つに大別することができる。

3.1 特徴点の抽出

スケール変化に対して不変な対応付けを行なうためには, スケールに応じた大きさの局所領域で特徴量を抽出する必要がある。このために, スケール空間でのフィルタリング処理を用いて特徴点を抽出する。SIFT では演算コスト低減のために式(2)の Difference of Gaussian (DoG) を用いて DoG 画像 $D(x, y, \sigma)$ を生成し, 特徴点の候補として $D(x, y, \sigma)$ からスケールと位置に対して極大値を示す画素を抽出する。

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (2)$$

ここで, $G(x, y, \sigma)$ は Gaussian フィルタの演算子, $I(x, y)$ は入力画像, σ はスケール, k はスケールの増加係数を表している。得られた特徴点の候補から, エッジ上の特徴点やコントラストが低い特徴点を除外することで最終的な特徴点 $(x_n, y_n, \sigma_n) \in N$ を抽出する。また, 画像の回転に対して不変な特徴量とするため, 各特徴点における勾配(方向)を算出する。具体的には, 特徴点を中心とした局所画像から得られる輝度勾配の強度と方向を用いてヒストグラムを作成し, ヒストグラム中の最大値と, 最大値の80%以上の値をとる方向を特徴点の方向とする。

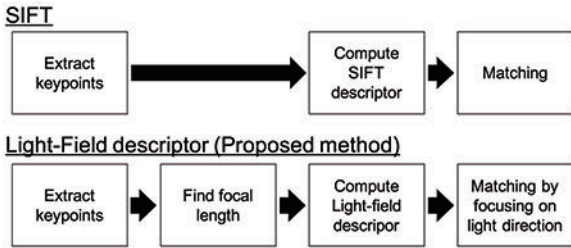


Fig. 3 Overview of SIFT and light-field descriptor (Proposed method)

3.2 特徴量の算出

SIFT 特徴量は、特徴点を中心とした局所画像から抽出する。また局所画像は特徴点のスケール σ に基づく Gaussian 関数で重み付けされる。重み付けされた局所画像を小領域に分割し、各小領域で特徴点の方向を考慮した輝度勾配のヒストグラムを作成する。これを SIFT 特徴量 $\mathbf{h}(I, x, y, \sigma)$ とする。経験的に、局所画像を $4 \times 4 = 16$ 個の小領域に分割し、輝度勾配ヒストグラムのビン数を 8 とするのが良いとされており、結果として 128 次元の特徴ベクトルが得られる。

3.3 特徴点の対応付け

特徴点の対応付けは特徴量間の L_2 距離に基づく最近傍探索によって行なう。しかし、特徴量が似ている特徴点と誤って対応付けされる場合がある。これを防ぐために、式 (3) のように、最も小さい L_2 距離 n_{1st} と 2 番目に小さい L_2 距離 n_{2nd} の比がしきい値 T 以上の場合のみ対応付けを行なう。

$$\frac{n_{2nd}}{n_{1st}} > T \tag{3}$$

4. ライトフィールド特徴量とその対応付け

2章で述べた通り、ライトフィールドは様々な方向の光線情報から構成され、部分開口画像として様々な視点の画像を得ることができる。本論文ではこの性質に着目し、光線方向を含む特徴量へと従来の局所特徴量を拡張したライトフィールド特徴量を新たに定義する。以下では SIFT に対するライトフィールド特徴量の拡張について説明を行なうが、他の特徴量に対しても同様に拡張が可能である。図 3 に SIFT と提案手法の全体構成をそれぞれ示す。3章で述べたとおり SIFT は、(1) 特徴点の抽出、(2) SIFT 特徴量の算出、(3) 特徴点の対応付け、の3つの処理から構成される。これに対して、提案手法は図 1 に示したとおり、(1) 特徴点の抽出、(2) 特徴点毎に合焦(焦点が合う)する焦点距離の推定、(3) 抽出された特徴点と推定された視差を用いてライトフィールド特徴量を算出、(4) 光線方向を考慮した対応付け、の処理から構成される。(1) の特徴点抽出ではライトフィールドから得られる中心画像 $I_{0,0}$ を入力画像とし、従来手法と同様に特徴点を抽出する。以降で、従来手法からの拡張について説明する、具体的には、まずライトフィールド特徴量について詳細に説明した後、光線方向を考慮した対応付けについて述べる。最後に特徴点毎で合焦する焦点距離の推定について述べる。

4.1 ライトフィールド特徴量

ライトフィールド特徴量は、シーン中の同一点を撮影した画像の特徴量として、部分開口画像毎に求めた特徴ベクトルの集合として定義される。具体的には、部分開口画像毎の特徴量を特徴空間に投影した特徴ベクトルの集合として表現され、これ

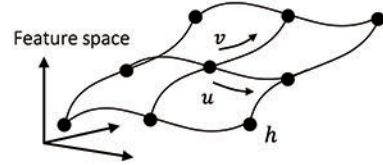


Fig. 4 Manifold in a feature space

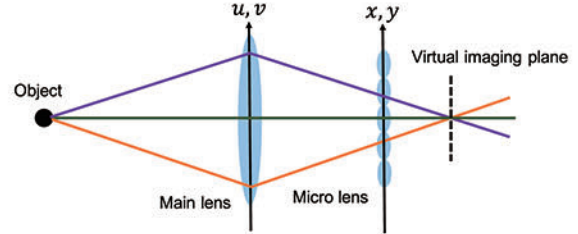


Fig. 5 Refocus mechanism

は光線方向を表す (u, v) の変化に応じて変化する特徴ベクトルの集合となる(図 4)。しかし、視差の影響により、異なる部分開口画像で、シーン中で同一の特徴点の座標は一致しないため、視差による座標の変化を考慮して特徴量を算出する必要がある。

ライトフィールドを用いたリフォーカスは、部分開口画像の視差を考慮して、部分開口画像を平行移動し、加算平均することで実現される。そこで、まずは部分開口画像を用いたリフォーカスについて説明する(図 5)。カメラの焦点距離を F とし、被写体に合焦する距離を αF とした場合、仮想撮像面を図 5 の位置に移動させることで、被写体に合焦するようにリフォーカスをすることができる。仮想撮像面の移動は、部分開口画像を u, v に応じて平行移動し、加算平均することで実現できて、リフォーカス後の画像 $I_\alpha(x, y)$ は式 (4) で定義される¹⁶⁾。

$$I_\alpha(x, y) = \sum_{u,v \in U,V} L\left(x + \left(1 - \frac{1}{\alpha}\right)u, y + \left(1 - \frac{1}{\alpha}\right)v, u, v\right) \tag{4}$$

ここでの平行移動量は、視差による部分開口画像毎の座標変化量と同じになるため、特徴点の部分開口画像毎の視差を考慮した座標は式 (5) のように定義することができる。

$$(x', y') = \left(x + \left(1 - \frac{1}{\alpha}\right)u, y + \left(1 - \frac{1}{\alpha}\right)v\right) \tag{5}$$

実際に視差を考慮した特徴点の座標を算出するには、事前に α を推定する必要がある。 α を推定する手法の詳細は 4.3 節で述べる。式 (5) を用いることで部分開口画像の特徴ベクトル \mathbf{h}_l は、光線方向を表す (u, v) と部分開口画像 $I_{u,v}$ を用いて式 (6) のように表される。

$$\mathbf{h}_l(L, x, y, u, v, \alpha, \sigma) = \mathbf{h}\left(I_{u,v}, x + \left(1 - \frac{1}{\alpha}\right)u, y + \left(1 - \frac{1}{\alpha}\right)v, \sigma\right) \tag{6}$$

ライトフィールド特徴量 \mathbf{H}_l は部分開口画像の特徴ベクトル \mathbf{h}_l の集合として表現され、次式で与えられる。

$$\mathbf{H}_l(L, x, y, \alpha, \sigma) = \{\mathbf{h}_l(L, x, y, u, v, \alpha, \sigma) | u \in U, v \in V\} \tag{7}$$

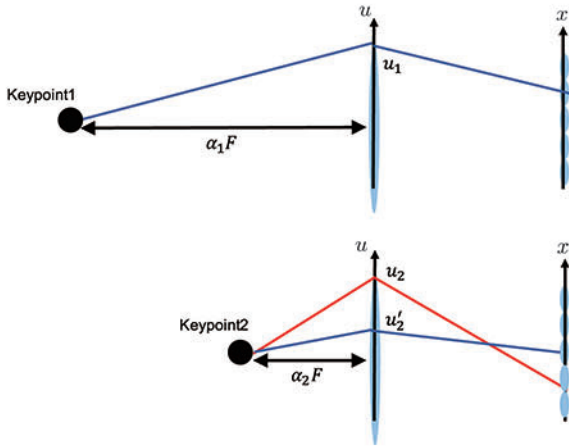


Fig. 6 Matching with the same direction of light ray

4.2 ライトフィールド特徴量の対応付け

特徴点の対応付けは、SIFTの場合と同様に、特徴ベクトルの L_2 距離に基づく最近傍探索によって行なう。ライトフィールド特徴量は、特徴点の位置 (x, y) に対して、多様な (u, v) に関する特徴ベクトルの集合で表現されるため、光線方向の角度（視線方向）が同一となる (u, v) における特徴ベクトルの L_2 距離で類似度を計算する必要がある。ここでは2つの特徴点 $(x_1, y_1, \sigma_1), (x_2, y_2, \sigma_2)$ のライトフィールド特徴量から類似度を算出する手法について述べる。合焦する焦点距離が異なる特徴点は、同じ (u, v) の部分開口画像であっても光線方向が異なる。光線方向が異なれば特徴量も異なるので、光線方向の角度を揃えて L_2 距離を計算する必要がある。図6は焦点距離が異なる特徴点からの光線を模式的に表している。なお、簡略化のため、 (x, u) 方向のみを示している。光線の方向を揃えて対応付けを行なうには特徴点1の u_1 に対して、光線の方向が等しくなるような特徴点2の u'_2 を求めて対応付けを行なう。相似の関係から、 u_1 から u'_2 を求める変換式は、式(8)で与えられる。

$$u'_2 = f(u) = \frac{\alpha_2}{\alpha_1} u_1 \quad (8)$$

$\alpha_1 F$ と $\alpha_2 F$ は、それぞれ特徴点1と特徴点2に合焦する焦点距離を表す。 v'_2 についても同様に求めることができ、全ての光線方向で計算した L_2 距離の平均をライトフィールド特徴量間の L_2 距離として定義する(式(9))。

$$n = \|\mathbf{H}_1(L_1, x_1, y_1, \alpha_1, \sigma_1) - \mathbf{H}_1(L_2, x_2, y_2, \alpha_2, \sigma_2)\| \\ = \frac{1}{N_{uv}} \sum_{u,v \in U, V} \|\mathbf{h}_1(L_1, x_1, y_1, u_1, v_1, \alpha_1, \sigma_1) \\ - \mathbf{h}_1(L_2, x_2, y_2, u'_2, v'_2, \alpha_2, \sigma_2)\| \quad (9)$$

ここで、 N_{uv} はライトフィールドから得られる u, v の総数を表している。観測における $u \in U, v \in V$ は離散的な値であるため、 u', v' の特徴量 $\mathbf{h}_1(L, x, y, u', v', \alpha, \sigma)$ は、図7に示すように補間する。従来手法と同様、最も小さい L_2 距離 n_{1st} と2番目に小さい L_2 距離 n_{2nd} の比がしきい値 T 以上の場合、その特徴点同士を対応付ける。

4.3 視差の推定

中心画像 $I_{0,0}$ から抽出した特徴点からライトフィールド特徴量を抽出し、対応付けを行なうためには、式(5)と式(8)にある各特徴点に合焦する焦点距離 α を推定する必要がある。ここでは簡略化のため、 $L(x, y, u, v)$ を x 方向の1軸に限定

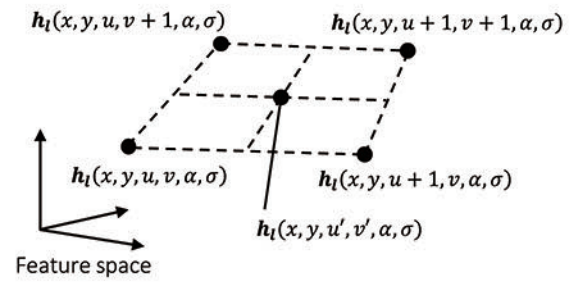


Fig. 7 Interpolation of descriptors

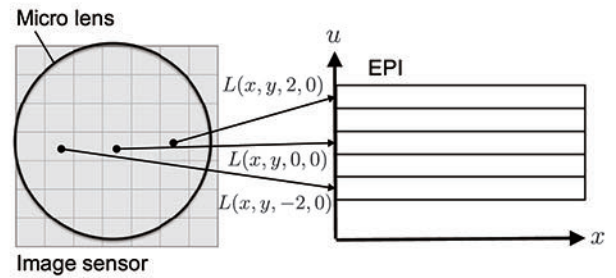


Fig. 8 Epipolar image

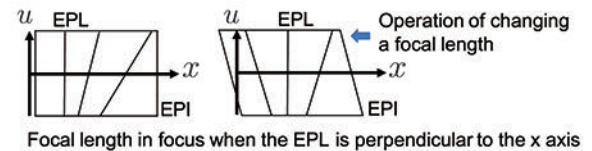


Fig. 9 Estimation of disparity from the epipolar line

した $L(x, u)$ を用いて焦点距離の推定を説明する。ここで、横軸 x 、縦軸 u とした部分開口画像の部分画像 (x, u) を考える(図8)。本論文では従来のライトフィールドに関連する論文の用語に倣い、このような画像を EpiPolar Image (EPI) と呼ぶ。被写体の同一点から各部分開口を通過する光はそれぞれの部分開口画像に記録される。これらの座標を EPI 中で結んだ線を EpiPolar Line (EPL) と呼ぶ。焦点距離 α を変化させると、図9に示すように EPI と EPL が変化する。注目画素で EPL の角度が x 軸に対して垂直でない場合は、部分開口画像の加算によりボケた画像が生成される。従って、合焦している場合は、EPL は x 軸に対して垂直な直線となる。次に、 $u = 0$ における部分開口画像とそれ以外の部分開口画像の差 SSD (Sum of Squared Difference) $S(\alpha, x)$ を考え、次式のように定式化する。

$$S(\alpha, x) = \frac{1}{|W_D|} \sum_{x \in W_D} \sum_u \left(L\left(x + \left(1 - \frac{1}{\alpha}\right)u, u\right) - L(x, 0) \right)^2 \quad (10)$$

ここで、 W_D は注目画素を中心とした一定の大きさの窓領域を表す。SSD は、 x 軸に対して EPL が垂直になったときに最小値をとり、垂直からの傾きが大きくなるほど SSD は大きくなる。従って、EPL と x 軸の角度が垂直になる、すなわち、SSD が最小となる α が、注目画素で合焦する焦点距離 $\hat{\alpha}$ である。そこで次式によって各特徴点に合焦する焦点距離 $\hat{\alpha}$ を求める。

$$\hat{\alpha} = \arg \min_{\alpha} S(\alpha, x) \quad (11)$$

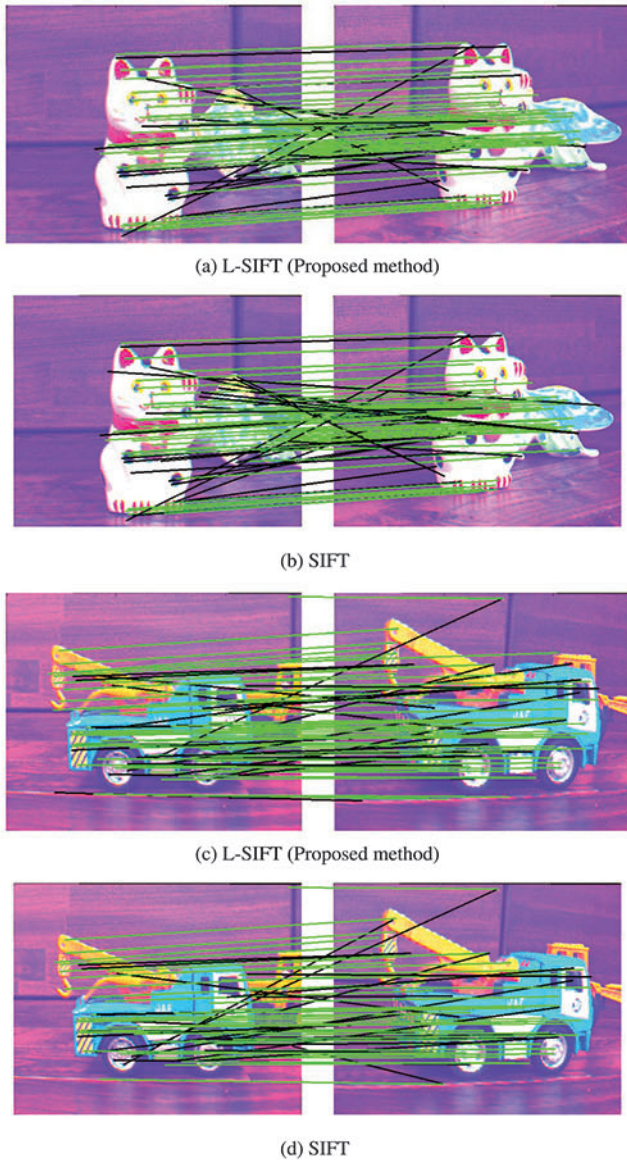


Fig. 10 Keypoint matching results of the LSFIT and the SIFT (Ratio of L_2 -distance larger than 1.2) on raw data images. Green line: correct matches, Black line: wrong matches

5. 実験

ここではライトフィールド特徴量による特徴点の対応付け精度を評価する方法と評価結果を述べる。本論文で提案するライトフィールド特徴量は、視線方向を一致させた対応付けを行なうことが可能で、視点変化が伴う画像間で特徴点の対応付け精度の向上を狙いとしている。また、被写体が立体物であっても対応付けの精度が低下しないという利点がある。ここでは、立体物に対して視点変化が伴う画像間の特徴点対応付け性能を評価する。次に一般的なシーンでの対応付け性能を評価するため、様々なシーンで撮影したデータを用いて対応付け性能の評価を行なう。ライトフィールド画像の取得には Lytro 社¹¹⁾のライトフィールドカメラ ILLUM を用いた。

5.1 立体物に対する対応付け性能の評価

立体的な被写体を複数視点から撮影したデータで評価を行ない、提案手法の効果を確認した。カメラ位置を変えて招き猫の人形とトラックの模型を撮影したライトフィールド画像をデー

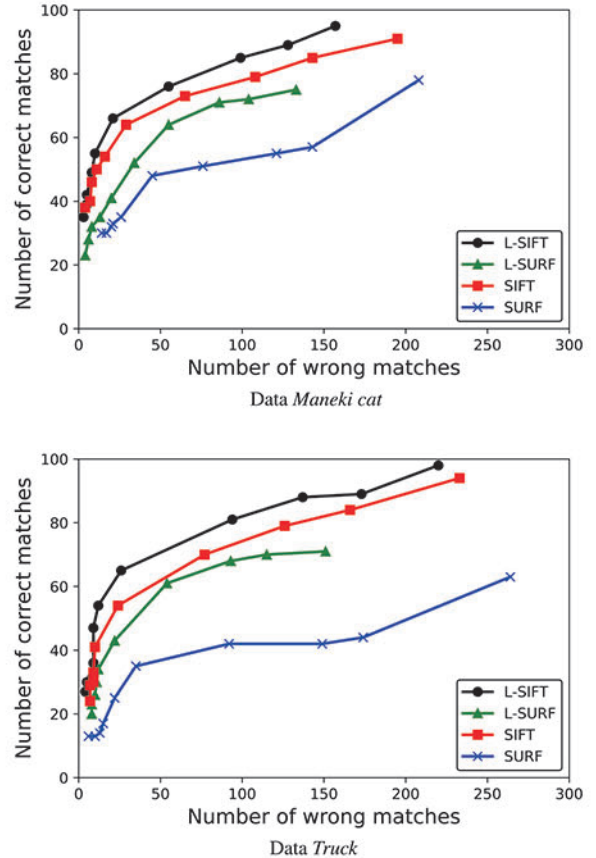


Fig. 11 Number of matches by L-SIFT, LSURF, SIFT, and SURF

タとした。4.1 節では SIFT に基づいてライトフィールド特徴量の詳細を述べたが、他の特徴量に対しても同様に拡張可能である。ここでは以下の手法の対応付け性能を比較した。

- L-SIFT (提案)：SIFT に基づくライトフィールド特徴量
- L-SURF (提案)：SURF に基づくライトフィールド特徴量
- SIFT：中心画像 $I_{0,0}$ から抽出した SIFT⁴⁾
- SURF：中心画像 $I_{0,0}$ から抽出した SURF¹⁴⁾

ただし、特徴量とその対応付けの比較に着目するため、L-SURF と SURF で用いる特徴点は、SIFT と同じ手法で抽出した。図 10 に Lytro ILLUM で撮影したライトフィールドから得られる中心画像 $I_{0,0}$ とライトフィールド SIFT 特徴量と SIFT 特徴量で対応付けした結果を示す。緑の線は正対応、黒の線が誤対応を示しており、式 (3) のしきい値 T を 1.2 とした場合の対応付け結果を表している。図 11 には式 (3) のしきい値を変化させた場合の正対応付け数と誤対応付け数をプロットした。正誤については人の目で判断した。図 11 から、基となる特徴量に関わらず、提案手法により正対応数が増加し、誤対応数が減少することを確認できた。同じしきい値の場合は特に誤対応数の低減に大きな効果が見られた。これらの評価結果からライトフィールド特徴量による特徴点対応付け精度の向上が確認できた。

5.2 様々なデータでの評価

一般的なシーンでの特徴点の対応付け性能を評価するため、実環境で撮影した 20 組のライトフィールド画像からなるデータで評価した。評価結果と合わせて、評価に用いたデータを

Table 1 Number of matches by L-SIFT and SIFT (Top30 L_2 -distance, Ratio of L_2 -distance larger than 1.2)

Dataset	Proposed method		SIFT	
	Correct	Wrong	Correct	Wrong
Boxes	29	1	26	4
Book (Rotation)	30	0	30	0
Appetizer	11	19	9	21
Bottle	13	17	12	18
Can	25	5	22	8
Pizza	23	7	17	13
Golf ball	8	22	4	26
Pole	28	2	25	5
Frog	22	8	22	8
Lion	11	19	8	22

Dataset	Proposed method		SIFT	
	Correct	Wrong	Correct	Wrong
Snack	20	10	15	15
Wall paint (Affine)	26	4	26	4
Dessert	19	11	19	11
Flag	26	4	26	4
Pasta 1	10	20	8	22
Sushi	23	7	22	8
Pasta 2	21	9	18	12
Shoes	24	6	23	7
Strawberiees	30	0	26	4
Gold fish	28	2	27	3

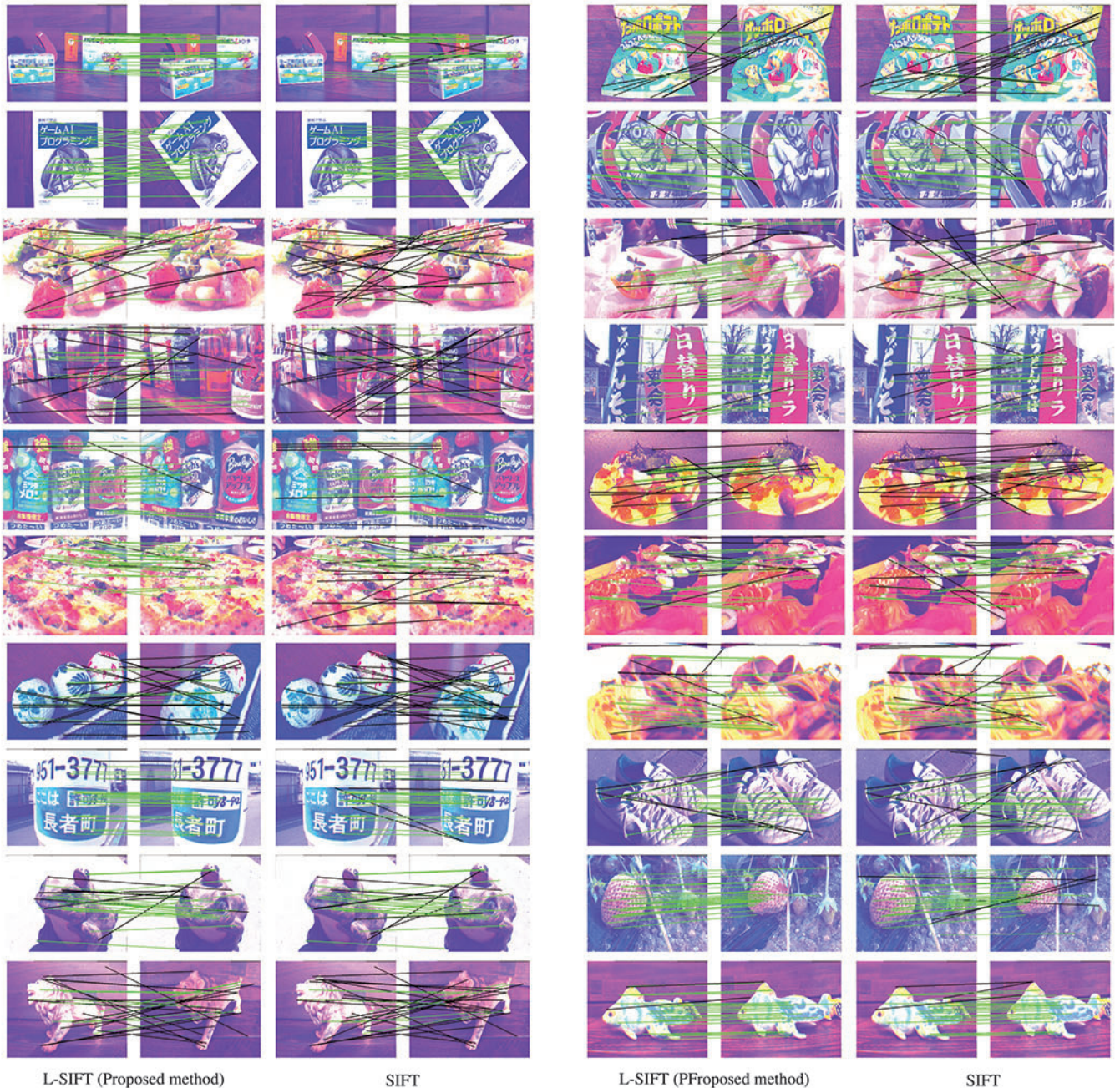


Fig. 12 Keypoint matchig results of L-SIFT and SIFT (Top30 L_2 -distance, Ratio of L_2 -distance larger than 1.2). Green line: correct matches, Black line: wrong matches

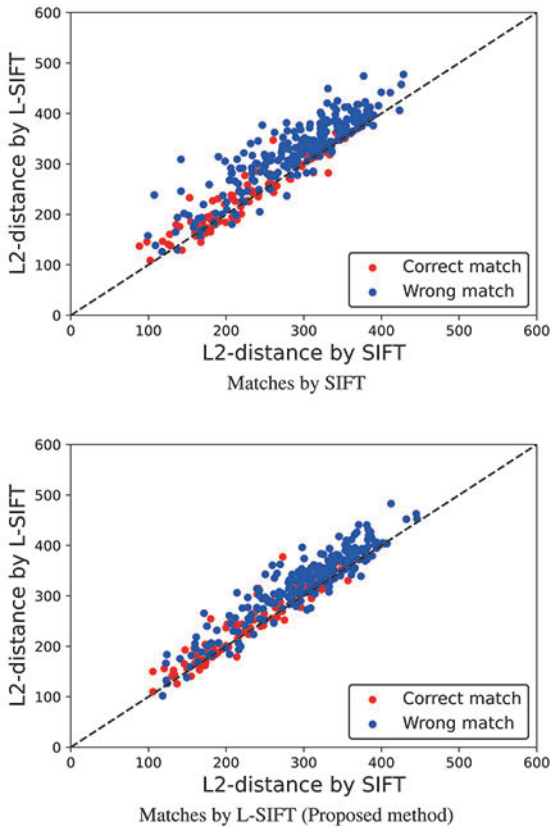


Fig. 13 L₂-distance of L-SIFT (Proposed method) and SIFT

図 12 に示す。L-SIFT (提案) と SIFT それぞれで対応付け性能を評価した。比較は式 (3) のしきい値 T を 1.2 とした時の対応付け結果のうち、 L_2 距離が小さい順に 30 組を抽出して正対応数と誤対応数により性能を評価した。表 1 に結果を示す。太字の数字は精度が高い手法を表している。評価の結果、6 つのデータを除いて、提案手法は SIFT より対応付け性能が向上することを確認した。6 つのデータ、具体的には Book (Rotation), Wall paint (Affine), Dessert, Flag, Frog では SIFT と同じ性能であった。Book (Rotation) は平面に対する回転操作のデータのため、SIFT においても不変な変換であることから、提案手法と差がなかったと考えられる。Wall paint (Affine), Dessert, Flag, Frog のデータについては、図 6 から分かる通り、合焦する焦点距離が遠い場合、ライトフィールドとして取得できる光線方向の角度範囲は小さくなる。そのため、通常のカメラから得られる画像データに対して、ライトフィールド特徴量の効果が表れにくく、SIFT と同等の対応付け精度になったと考えられる。焦点距離の違いが対応付け性能へ及ぼす影響については考察で詳細に述べる。

6. 考 察

6.1 ライトフィールド特徴量と SIFT 特徴量から算出した L_2 距離の比較

5 章で、ライトフィールド特徴量による特徴点の対応付け精度の評価を行なった。ここでは、L-SIFT と SIFT で計算した L_2 距離を比較し、対応付け精度に対するライトフィールド特徴量の効果について考察する。図 13 は対応付けられた特徴点の SIFT 特徴量の L_2 距離と L-SIFT の L_2 距離を图示したグラ

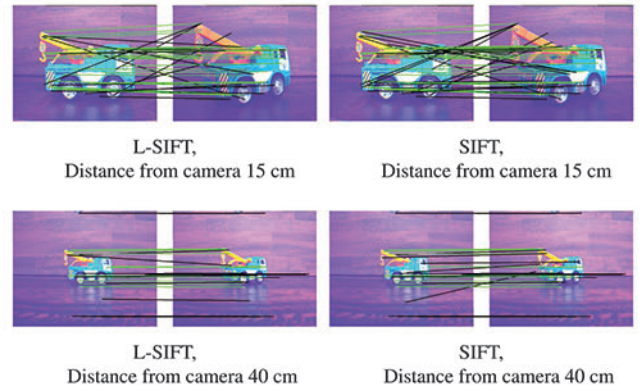


Fig. 14 Results of matching by L-SIFT (Proposed method) and SIFT with different distances from camera, Ratio of L_2 -distance larger than 1.2. Green line: correct matches, Black line: wrong matches

フである。横軸は SIFT 特徴量の L_2 距離、縦軸は L-SIFT の L_2 距離、赤い点は正対応、青い点は誤対応を示している。グラフ中の破線は L-SIFT と SIFT の L_2 距離が等しいことを表す。破線より左上の領域は L-SIFT の L_2 距離の方が SIFT の L_2 距離より大きい、逆に破線より右下の領域は SIFT の L_2 距離の方が L-SIFT の L_2 距離に比べて小さい領域を表す。誤対応を示す青い点が破線より左上の領域にあれば、L-SIFT の L_2 距離は SIFT よりも大きく、対応付けの判定精度が向上していると言える。SIFT で対応付けした場合のグラフを見ると破線の左上の領域に見られた誤対応の点があるが、L-SIFT で対応付けを行なった結果を見ると、破線左上の領域にあった誤対応の点が減少しているのがわかる。すなわちライトフィールド特徴量により、視点変化による特徴量の変化が特徴点間の L_2 距離に反映されていると言える。これらの結果、提案手法は立体的な被写体の視点変化による特徴量の変化が特徴点間の L_2 距離に反映され、より高精度な対応付けが可能になったことを確認できる。

6.2 被写体とカメラの距離に応じた対応付け性能の比較

提案手法の効果が表われるのは、ライトフィールドに広い角度の光線方向の情報が含まれている場合である。図 6 から分かる通り、ライトフィールドカメラはその構造上、被写体に近い距離で撮影した方がより広角な光線情報を取得することができる。従ってカメラと被写体の距離が対応付け性能に影響を及ぼすと考えられる。そこで以下の 2 つのデータを用いて評価を行ない、影響を検証する。

- 広角データ：広い角度の光線情報、被写体までの距離が 15cm のライトフィールド画像
- 狭角データ：狭い角度の光線情報、被写体までの距離が 40cm のライトフィールド画像

図 14 にそれぞれのデータで対応付けを行なった結果を示す。図 15 はしきい値毎の正対応付け数と誤対応付け数を示したグラフである。図 15 から、狭角データでは SIFT に対して、L-SIFT の性能向上幅が小さくなることが確認できる。

更に、6.1 節と同様に、SIFT 特徴量の L_2 距離と L-SIFT の L_2 距離の比較を行なった。ここでは L_2 距離が小さい順に 30 組を抽出して評価を行なった。SIFT で対応付けを行なった結果を図 16 に示す。図 16 (a) は広角データ、図 16 (b) は狭角データの結果である。広角データの対応付け結果では、左上

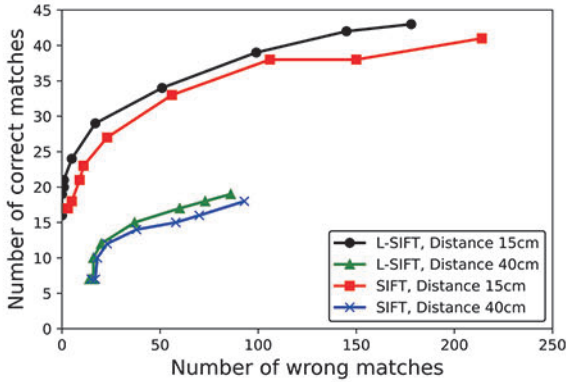


Fig. 15 Number of matches of L-SIFT (Proposed method) and SIFT with different distances from camera, Top30 L_2 -distance, Ratio of L_2 -distance larger than 1.2

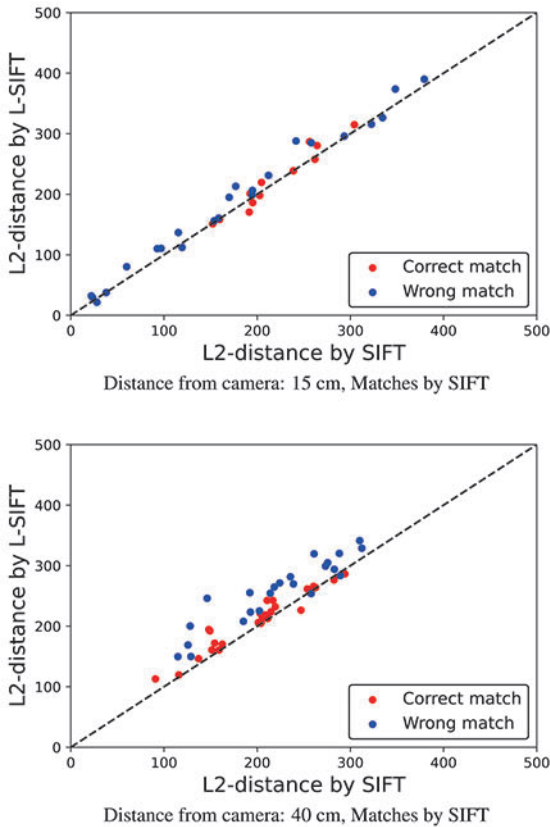


Fig. 16 Result of L_2 -Norm by light-field descriptor and SIFT with different distances from camera, Ratio of L_2 -distance larger than 1.2

の領域に誤対応の点が多く見られる。このことは、誤対応した特徴点の L-SIFT の L_2 距離が SIFT より大きくなり、L-SIFT で対応付け精度が向上することを示している。一方、狭角データでは全ての点が破線に近い。これは、L-SIFT と SIFT で L_2 距離がほぼ等しく、対応付け精度の性能向上の効果が小さくなることを示している。以上の結果、被写体までの距離が対応付け性能の結果に影響を及ぼすことを確認した。被写体までの距離が近い場合、ライトフィールドから得られる部分開口画像の視差が大きくなる。ライトフィールド特徴量は部分開口画像から得られる特徴量の集合で、視差の大きい部分開口画像であ

ば、特徴量の表現できる視差の範囲が大きくなり、対応付け性能が向上したと考える。現状のライトフィールドカメラで視差の大きな部分開口画像を得るには被写体に近づく必要がある。今後、被写体までの距離に関わらず視差の大きな部分開口画像を得られるようなライトフィールドカメラの研究を進めていく必要がある。

7. 結 言

本論文では、ライトフィールドから光線方向を考慮したライトフィールド特徴量を提案した。また、ライトフィールド特徴量による対応付けの方法を提案した。ライトフィールド特徴量を抽出して対応付けを行なうためには、特徴点毎の合焦する焦点距離が必要になり、その推定手法も合わせて提案した。提案手法の有効性を確認するため、立体物に対して複数視点のデータと実環境において撮影した 20 シーンの評価画像を用いて精度の検証を行なった。その結果、従来手法に比べ提案手法の対応付け性能が向上することを確認した。特に立体物に対して視点が増えたデータの対応付け性能向上を確認できた。また、ライトフィールドカメラの構造から記録できる光線の角度は被写体とカメラからの距離に応じて変化することを述べ、実データで焦点距離が近くなるほど提案手法の効果が高くなることを示した。ライトフィールド特徴量は部分開口画像の視差が大きい程、特徴量の表現できる視差の範囲が大きくなる。これらの特性から現状のライトフィールドカメラでは、例えば、顕微鏡の画像、回路パターンなどの画像、医療カメラの画像など、接写できるシーンでの応用が期待できる。今後、被写体との距離によらず対応付けの性能を向上するには、大きなダイナミックレンジで光線情報が取得できるようなライトフィールドカメラの研究を進めていく必要がある。また更なる性能向上に向けては、ライトフィールドを活用した特徴点の抽出、ライトフィールド特徴量の演算と対応付けの演算コストを削減する手法の検討を進めていく必要がある。ライトフィールド特徴量は特徴量の集合になるため、特徴量が大きくなり、その算出や対応付けで演算コストが高くなる。特に部分開口画像の視差が小さい場合は、似た特徴量の集合になってしまう。従って、PCA-SIFT⁵⁾のような主成分分析を用いた特徴量の次元削減は、特徴量の情報量を減らすことなく演算コストを削減することに有効と考える。今回の研究ではマイクロレンズのばらつきやノイズを考慮せず、理想的な状態で光線の方向を算出して対応付けを行なっている。これらを考慮した対応付けの手法や校正手法の検討も必要と考える。

謝 辞

本研究の一部は、科学研究費補助金 JP17H00745 による。

参 考 文 献

- 1) M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, **74**, 1, (2007) 59.
- 2) G. Csurka, C.R. Dance, L. Fan, J. Willamowski and C. Bray, "Visual categorization with bags of keypoints," *Proc. ECCV2004 Workshop on Statistical Learning in Computer Vision*, (2004) 59.
- 3) R. Mur-Artal, J. M. M. Montiel and J.D. Tardes, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, **31**, 5, (2015) 1147.
- 4) D. Lowe, "Distinctive image features from scaleinvariant keypoints," *Int. J. Comput. Vis.*, **60**, 2, (2004) 91.
- 5) Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," *Proc. 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, (2004) 506.
- 6) E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An effi-

- cient alternative to SIFT or SURF,” Proc. 13th Int. Conf. on Computer Vision, (2011) 2564.
- 7) A. Alahi, R. Ortiz and P. Vandergheynst, “FREAK: Fast retina keypoint,” Proc. 2012 IEEE Conf. on Computer Vision and Pattern Recognition, (2012) 510.
 - 8) J. Matas, O. Chum, M. Urban and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” Image Vision Comput., **22**, 10, (2004) 761.
 - 9) K. Mikolajczyk and C. Schmid, “Scale & affine invariant interest point detectors,” Int. J. Comput. Vis., **60**, 1, (2004) 63.
 - 10) V. Vaish, B. Wilburn, N. Joshi and M. Levoy, “Using plane + parallax for calibrating dense camera arrays,” Proc. 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, (2004) 2.
 - 11) “Lytro,” <http://www.lytro.com/>
 - 12) “raytric,” <https://raytrix.de/>
 - 13) 蚊野浩, “ライトフィールドカメラ Lytro の動作原理とアルゴリズム,” 応用物理学会第 127 回微小光学研究会, 4, (2013).
 - 14) H. Bay, T. Tuytelaars and L. Van Gool, “SURF: Speeded up robust features,” Proc. 9th European Conf. on Computer Vision, 1, (2006) 404.
 - 15) J.-M. Morel and G. Yu, “ASIFT: A new framework for fully affine invariant image comparison,” SIAM J. Imaging Sci., **2**, 2, (2009) 438.
 - 16) R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz and P. Hanrahan, “Light field photography with a hand-held plenoptic camera,” Stanford University Computer Science Tech. Rep., CSTR 2005-02, (2005).
 - 17) R. Ng, “Digital light field photography,” Ph. D thesis, Stanford University, (2006).
 - 18) M. W. Tao, S. Hadap, J. Malik and R. Ramamoorthi, “Depth from combining defocus and correspondence using light-field cameras,” Proc. 14th IEEE Int. Conf. on Computer Vision, (2013) 673.
 - 19) 清水政行, 川西康友, 出口大輔, 井手一郎, 村瀬洋, “ライトフィールド情報を活用した特徴点の対応付けに関する検討,” 電子情報通信学会技術研究報告, PRMU2017-36, (2017).
 - 20) J. S. Lumentut, Williem and I. K. Park, “Feature descriptors for 4D light field images,” Proc. 23rd Japan-Korea Workshop on Frontiers of Computer Vision, O3-5, (2017).
 - 21) I. Tomic and K. Berkner, “3D keypoint detection by light field scale-depth space analysis,” Proc. 2014 IEEE Int. Conf. on Image Processing, (2014) 1927.