

多視点ステレオのための 位相限定相関法に基づく画像マッチングの高精度化

酒井 修二¹ 伊藤 康一¹ 青木 孝文¹ 増田 智仁² 運天 弘樹²

概要: 多視点ステレオアルゴリズムにおいて、3次元点を推定するための画像マッチング手法は、アルゴリズムの精度、ロバスト性、計算コストに大きく影響する重要な要素である。これまでに、筆者らは、位相限定相関法 (Phase-Only Correlation: POC) を多視点画像間のマッチングに用いることで、従来の正規化相互相関 (Normalized Cross-Correlation: NCC) と比べて、少ない計算コストで高精度な3次元復元が実現できることを示した。本稿では、POCに基づく画像マッチング手法を高精度化するために、対象物体の形状を考慮したマッチングを適用し、ステレオ画像間の変形を軽減する手法を提案する。実験を通して、基線長が長い場合などのようにステレオ画像間の変形が大きい場合に3次元復元精度が向上することを示す。

1. はじめに

多視点ステレオアルゴリズムは、異なる視点から撮影された複数枚のカメラ画像を用いて物体の3次元形状を復元する技術である [1], [2], [3]。近年、計算機やカメラの技術的發展や、3次元復元技術の進歩に伴い、カメラ画像のみを用いて高品質な3次元形状モデルが生成できるようになった。これにより、多視点ステレオアルゴリズムは、産業、医療、芸術など、様々な分野で注目されるようになった。多視点ステレオアルゴリズムは、(i) 画像マッチングを用いた3次元点群の生成、(ii) 3次元点群からの3次元モデル生成といった処理で構成される。これらの処理の中でも、画像マッチングは、多視点ステレオアルゴリズム全体の精度、ロバスト性、計算コストに大きく影響し、アルゴリズムの性能を決定する重要な要素である。

多くの多視点ステレオアルゴリズムでは、正規化相互相関 (Normalized Cross-Correlation: NCC) に基づく画像マッチングが用いられている [1], [4], [5], [6], [7], [8], [9]。Goesele らは、NCC に基づく画像マッチングをプレーンスweepingに適用し、奥行きを変化させながら複数のステレオ画像ペアで計算される相関値を足し合わせることで、信頼性の高い奥行きマップ生成を実現した [4]。Campbell らは、Goesele らの手法に対して、周辺画素のマッチング

結果を利用することで、高精度かつ誤対応の少ない奥行きマップ生成を行っている [7]。さらに、Bradley らや、Furukawa らは、3次元点の奥行きだけでなく法線ベクトルの向きも考慮して、マッチングウィンドウを変形させることで、よりロバストなマッチングを行った [6], [9]。

これらの従来手法では、NCC に基づく画像マッチングにより、ある仮定された3次元点かどの程度尤もらしいかを判定する。そのため、3次元点のパラメータ（奥行きや法線ベクトル）を変化させながらマッチングを繰り返し行い、最も尤もらしいパラメータを探索する必要がある。例えば、Goesele らの手法に代表されるプレーンスweepingでは、3次元点の奥行きを離散的に変化させながらNCCのマッチングスコアを計算し、最もスコアが高くなる3次元点の奥行きを採択する。このとき、高精度な3次元復元を行うためには、奥行き変化の刻み幅を非常に細かくする必要があり、マッチング回数が膨大になるという問題があった。また、奥行き変化の刻み幅が非常に小さい場合、ステレオ画像上では3次元点の移動幅がサブピクセルとなる。多くの多視点ステレオアルゴリズムでは、サブピクセルの輝度値を線形補間で求めているが、自然画像においてピクセル間の輝度値が線形に変化しているという仮定は必ずしも成立するわけではない。

これらの問題に対して、筆者らは、多視点ステレオのための位相限定相関法 (Phase-Only Correlation: POC) に基づく画像マッチング手法を提案している [10]。POC は、画像の位相情報に着目した画像マッチング手法であり、解析的なピークモデルを用いた高精度な平行移動量推定が可能である [11]。この POC を多視点ステレオ画像間の局所的

¹ 東北大学 大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University,
Sendai-shi, Miyagi, 980-8579, Japan

² 凸版印刷株式会社
Toppan Printing Co., Ltd., Bunkyo-ku, Tokyo, 112-8531,
Japan

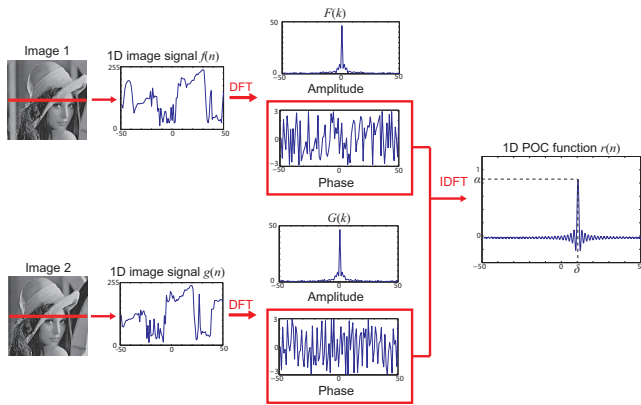


図 1 1次元 POC に基づく画像マッチング

なウィンドウマッチングに適用し、ステレオ画像間の局所的な平行移動量から 3 次元座標を推定することで、従来の NCC に基づく画像マッチングと比べて非常に少ない回数の画像マッチングで高精度な 3 次元復元を行うことが可能である。しかしながら、POC では画像信号間の変形を平行移動のみと仮定するため、ステレオ画像間の基線長が長い場合や、3 次元復元対象がカメラに対して大きく傾いている場合など、ステレオ画像間の変形が大きい場合に 3 次元復元誤差が大きくなるという問題があった。そこで、本稿では、3 次元復元対象の表面形状を考慮した画像変形の軽減を行うことで、多視点ステレオのための POC に基づく画像マッチングの高精度化を行う。また、Strecha らにより公開されているデータセット [12] を用いた性能評価実験を通して、提案手法による 3 次元復元精度の向上を確認する。

2. 多視点ステレオのための位相限定相関法に基づく画像マッチング

本章では、多視点ステレオのための POC に基づく画像マッチング [10] について述べる。まず、POC の基本原理について概説し、次に、多視点ステレオのための POC に基づく画像マッチングについて説明する。

2.1 位相限定相関法

POC は、画像の位相成分に着目した画像マッチング手法である [11]。平行化後のステレオ画像ペアのように、視差が 1 次元方向に限定される場合、1 次元の POC 関数が有効である [13]。以下では、 N 点の 2 つの 1 次元画像信号 $f(n)$ および $g(n)$ が与えられたとき、1 次元 POC を用いて 2 つの画像信号間の類似度と平行移動量を算出する方法について述べる。

図 1 に 1 次元 POC の計算の流れを示す。1 次元画像信号 $f(n)$ および $g(n)$ の 1 次元離散フーリエ変換 (Discrete Fourier Transform: DFT) を $F(k)$, $G(k)$ とするとき、正規化相互パワースペクトル $R(k)$ を次式で定義する。

$$R(k) = \frac{F(k)\overline{G(k)}}{|F(k)\overline{G(k)}|} = e^{j(\theta_F(k) - \theta_G(k))} \quad (1)$$

ただし、 $\overline{G(k)}$ は $G(k)$ の複素共役であり、 $\theta_F(k) - \theta_G(k)$ は 2 つの 1 次元画像信号の位相差スペクトルである。また、 $k = -M, \dots, M$ は離散周波数 インデックスであり、 $N = 2M + 1$ である。このとき、1 次元 POC 関数 $r(n)$ は、正規化相互パワースペクトル $R(k)$ の 1 次元逆離散フーリエ変換 (Inverse Discrete Fourier Transform: IDFT) として以下の式で表される。

$$r(n) = \frac{1}{N} \sum_{k=-M}^M R(k) W_N^{-kn} \quad (2)$$

ここで、 $W_N = e^{-j\frac{2\pi}{N}}$ は回転因子である。1 次元画像信号 $f(n)$ と $g(n)$ が互いに微小量 δ だけ平行移動した関係にあると仮定すると、 $f(n)$ と $g(n)$ の POC 関数 $r(n)$ は次式で与えられる。

$$r(n) \simeq \frac{\alpha \sin(\pi(n + \delta))}{N \sin(\frac{\pi}{N}(n + \delta))} \quad (3)$$

上式は、1 次元画像信号が δ だけ微小に平行移動した場合の POC 関数のピーク形状を表している。このとき、POC 関数の相関ピークの高さ α は 2 つの信号の類似度の指標に、相関ピークの位置座標 δ は 2 つの信号の平行移動量に相当する。このパラメータ α と δ を推定することで、信号の類似度とサブピクセル精度の平行移動量を求めることができる。POC を用いた画像マッチングでは、計算される POC 関数の離散データに対して、式 (3) の解析的ピークモデルをフィッティングすることで、画像のピクセル間に存在するピーク座標を推定する。さらに、(i) 窓関数の適用、(ii) スペクトル重み付け、(iii) 複数信号の利用を適用することで、高精度な平行移動量推定が可能である [11], [13]。

2.2 多視点ステレオのための POC に基づく画像マッチング

POC を多視点ステレオに適用する場合、多視点画像から複数のステレオペアを作成し、POC を用いて各ステレオ画像上の局所的な平行移動量を求めることで、3 次元点の奥行きを高精度に推定する。このとき、画像上の平行移動量はステレオ画像ごとに異なるため、POC 関数のピーク座標もステレオ画像ごとに異なる。そこで、正規化視差の概念に基づき、各ステレオペアから計算される POC 関数のピーク座標が一致するようにマッチングウィンドウの大きさを決定することで、複数のステレオペアから計算される POC 関数を統合する [10]。以下では、参照視点における視線上の 3 次元点の初期値 $\mathbf{M}'$ から、POC を用いて真の 3 次元点 $\mathbf{M}$ を求める問題を考える。

図 2 に、3 次元点 1 点当たりの処理の流れを示す。まず、カメラパラメータが既知の多視点画像 $\mathbf{V} = \{V_0, \dots, V_{H-1}\}$

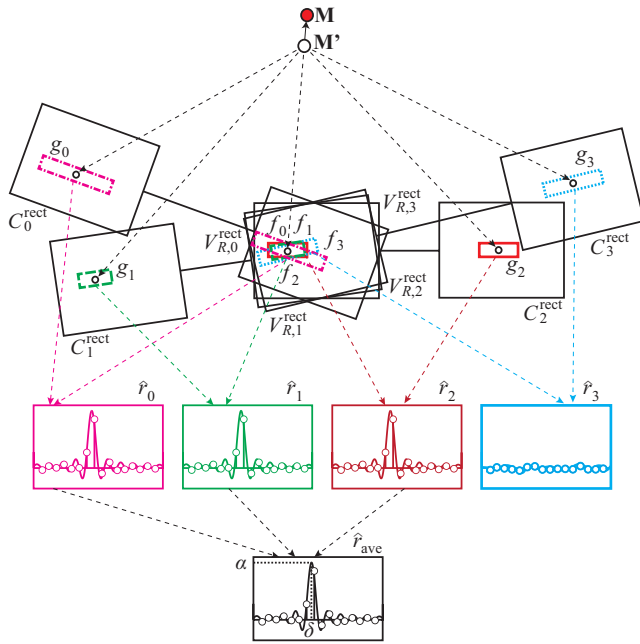


図 2 多視点画像間の POC による奥行き推定

のうち、参照視点を $V_R \in \mathbf{V}$ 、その近傍視点 $\mathbf{C} = \{C_0, \dots, C_{K-1}\} \subset \mathbf{V} - \{V_R\}$ とする。ここで、 H および K は、それぞれ多視点画像の数と近傍視点の数である。 V_R および $\mathbf{C}$ から K 組の平行ステレオペア $V_{R,i}^{\text{rect}} - C_i^{\text{rect}}$ ($i = 0, \dots, K-1$) を生成する [1]。3次元点の初期値 $\mathbf{M}'$ を平行ステレオペア $V_{R,i}^{\text{rect}} - C_i^{\text{rect}}$ に投影し、それぞれ画像座標を $\mathbf{m}_i = [u_i, v_i]$ および $\mathbf{m}_i^C = [u_i^C, v_i^C]$ とする。 $V_{R,i}^{\text{rect}}$ から $\mathbf{m}_i$ を中心に $s_i w \times L$ ピクセルの大きさでウィンドウ f_i を切り出し、 C_i^{rect} から $\mathbf{m}_i^C$ を中心に $s_i w \times L$ の大きさでウィンドウ g_i を切り出す。ここで、 $w \times L$ は基準となるマッチングウィンドウの幅と高さを表す。 s_i は視差の正規化のための拡大縮小率であり、多視点画像のカメラパラメータと 3次元点の参照視点上の画像座標から計算される [10]。切り出されたウィンドウ f_i, g_i を $w \times L$ の大きさに揃えるために $1/s_i$ 倍に拡大縮小し、拡大縮小されたウィンドウから 1次元 POC 関数 $\hat{r}_i$ を計算する。以上のように計算された 1次元 POC 関数 $\hat{r}_i$ ($i = 0, \dots, K-1$) では、POC 関数のピーク座標が一致する。このことより、複数の 1次元 POC 関数 $\hat{r}_i$ ($i = 0, \dots, K-1$) の平均である $\hat{r}_{\text{ave}}$ を用いて、平行移動量推定の精度を向上させる。このとき、POC 関数のピーク値 α_i が閾値 th_{corr} を超えるもののみを用いて $\hat{r}_{\text{ave}}$ を計算することで、オクルージョンや物体境界の影響などを抑える。 $\hat{r}_{\text{ave}}$ に対して関数フィッティングを行い、POC 関数のピーク値 α とサブピクセルレベルのピーク座標 δ を得る。この平行移動量 δ より、3次元点の初期値 $\mathbf{M}'$ と真の 3次元点 $\mathbf{M}$ との移動量を計算し、真の 3次元点 $\mathbf{M}$ の座標を得る。

以上で説明した多視点ステレオのための POC に基づく画像マッチングでは、3次元点の初期値 $\mathbf{M}'$ が真の 3次

元点 $\mathbf{M}$ にある程度近くなければ、3次元点 $\mathbf{M}$ の座標を求めることができない。そこで、多視点ステレオのための POC に基づく画像マッチングを、プレーンスライピングや画像ピラミッドを用いた階層的探索に適用することで、参照視点画像上の任意の点の 3次元点の奥行きを高精度に求めることが可能である [10]。

3. 画像変形の軽減による多視点ステレオのための POC の高精度化

2章で述べた多視点ステレオのための POC では、2つの画像信号間の変形を平行移動のみと仮定しているため、ステレオペア間の基線長が長い場合や、3次元復元対象がカメラに対して大きく傾いている場合のように、ステレオ画像間の変形が大きい場合に 3次元復元誤差が大きくなるという問題があった。そこで、本章では、ステレオ画像間の変形を軽減するようなマッチングウィンドウの設定方法を提案し、多視点ステレオのための POC を高精度化する。ステレオ画像間の変形は、撮影された 3次元復元対象の表面形状やカメラの位置関係に依存する非線形な変形である。提案手法では、ステレオ画像間の変形を局所的な平行移動と拡大縮小、および、スキューで表現し、復元対象の物体表面の法線ベクトル $\mathbf{n}$ から拡大縮小率とスキューの傾きを推定する。

2章で述べたように、多視点ステレオのための POC では、多視点画像から複数の平行ステレオペアを生成し、各平行ステレオペア間において POC に基づく画像マッチングを行う。このとき、平行ステレオ画像間では、エピポーラ線が垂直座標軸、もしくは、水平座標軸と平行になるため、その画像変形は 1次元方向に限定される。そのため、図 3 に示すように、復元対象が局所的に平面であると仮定すると、平行ステレオペアの参照視点画像 $V_{R,i}^{\text{rect}}$ と近傍視点画像 C_i^{rect} のそれぞれのマッチングウィンドウ間の画像変形は拡大縮小とスキューによって近似できる。以下では、平行ステレオペア $V_{R,i}^{\text{rect}} - C_i^{\text{rect}}$ において、3次元点 $\mathbf{M}_i = [X_i, Y_i, Z_i]$ と $\mathbf{M}_i$ の法線ベクトル $\mathbf{n}_i = [n_{X,i}, n_{Y,i}, n_{Z,i}]$ が与えられたとき、拡大縮小率 ξ_i とスキューの傾き κ_i を計算する方法について述べる。ここで、 $\mathbf{M}_i$ および $\mathbf{n}_i$ の座標系は平行化後の参照視点 $V_{R,i}^{\text{rect}}$ のカメラ座標系であるが、同じ参照視点でも、平行化の処理の際に、ペアとなる近傍視点のカメラパラメータによってカメラ座標が回転することに注意されたい。また、これらの画像変形の軽減を伴う POC による 3次元復元手法の処理について具体的に述べる。

3.1 拡大縮小の軽減

まず、図 3 に示すように、エピポーラ線が水平座標軸と平行になるように平行化した場合について考える。3次元点 $\mathbf{M}_i$ と法線ベクトル $\mathbf{n}_i$ により定義される 3次元平面とエピポーラ平面の交線上で、 $\mathbf{M}_i$ を中心とする単位線分を

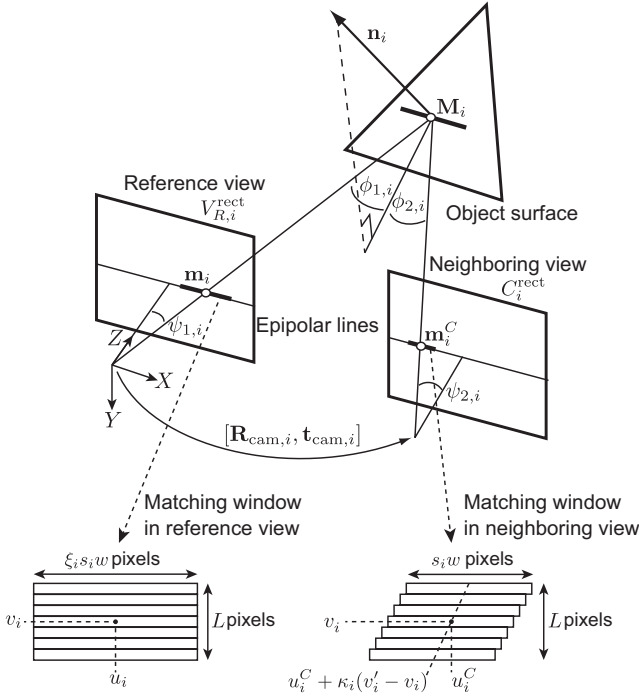


図 3 平行ステレオペアの左右の画像におけるマッチングウィンドウ

仮定する。この単位線分に対応する参照視点画像 $V_{R,i}^{rect}$ および近傍視点画像 C_i^{rect} 上の線分の長さ $w_{1,i}$, $w_{2,i}$ はそれぞれ次式で計算される [6]。

$$w_{1,i} = \frac{\cos \phi_{1,i}}{Z_i \cos \psi_{1,i}} \quad (4)$$

$$w_{2,i} = \frac{\cos \phi_{2,i}}{Z_i \cos \psi_{2,i}} \quad (5)$$

ここで、 $\psi_{1,i}$, $\psi_{2,i}$ はそれぞれ $V_{R,i}^{rect}$ および C_i^{rect} における光軸と視線のなす角を表し、 $\phi_{1,i}$, $\phi_{2,i}$ はそれぞれ $V_{R,i}^{rect}$ および C_i^{rect} における視線と $\mathbf{n}_i$ のエピポーラ平面への投影ベクトルのなす角を表す。式 (4) および式 (5) より、 $w_{1,i}$ と $w_{2,i}$ の関係は、

$$w_{1,i} = \frac{\cos \psi_{2,i} \cos \phi_{1,i}}{\cos \psi_{1,i} \cos \phi_{2,i}} w_{2,i} \quad (6)$$

$$= \xi_i w_{2,i} \quad (7)$$

となり、拡大縮小率 ξ_i は次式で表される [6], [14]。

$$\xi_i = \frac{\cos \psi_{2,i} \cos \phi_{1,i}}{\cos \psi_{1,i} \cos \phi_{2,i}} \quad (8)$$

また、エピポーラ線が垂直座標軸と平行になるように平行化した場合においても、同様に考えると、拡大縮小率は式 (8) で表される。実際の処理では、ウィンドウサイズ $w_{2,i}$ は単位線分の投影ではなく、2 章で説明した正規化視差を考慮したウィンドウサイズ $s_i w \times L$ となる。

以上より、提案手法では、参照視点におけるマッチングウィンドウを ξ_i 倍し、 $V_{R,i}^{rect}$ におけるウィンドウサイズを $\xi_i s_i w \times L$ ピクセル、 C_i^{rect} におけるウィンドウサイズを $s_i w \times L$ ピクセルとすることで、ステレオ画像間の局所的な拡大縮小を軽減する。ここで、参照視点のウィンドウを

ξ_i 倍するのではなく、近傍視点のウィンドウを $1/\xi_i$ 倍した場合、POC 関数 $\hat{r}_i$ のピーク座標が ξ_i 倍されるため、異なるステレオペア $V_{R,j}^{rect}-C_j^{rect}$ ($j \in \{1, \dots, K-1\} - \{i\}$) から計算される POC 関数 $\hat{r}_j$ と $\hat{r}_i$ のピーク座標が一致しないことに注意されたい。

3.2 スキューの軽減

まず、エピポーラ線が水平座標軸と平行になるように平行化した場合について考える。 $\mathbf{M}_i$ と $\mathbf{n}_i$ により定義される 3 次元平面上のある直線が、参照視点 $V_{R,i}^{rect}$ における $\mathbf{M}_i$ の対応点 $\mathbf{m}_i = [u_i, v_i]$ を通り垂直座標軸に平行な直線に投影されると仮定する。この 3 次元空間上の直線は、近傍視点 C_i^{rect} 上で $\mathbf{m}_i^C = [u_i^C, v_i^C]$ を通り、 κ_i の傾きを持つ直線として投影される。このとき、この 3 次元空間上の直線は、次式を満たす $[X, Y, Z]$ で表現される。

$$\begin{cases} n_{X,i}(X-X_i) + n_{Y,i}(Y-Y_i) + n_{Z,i}(Z-Z_i) = 0 \\ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} (u_i - u_{0i}) \|\mathbf{t}_{cam,i}\| / d_i \\ (v_i' - v_{0i}) \|\mathbf{t}_{cam,i}\| / d_i \\ \beta_i \|\mathbf{t}_{cam,i}\| / d_i \end{bmatrix} \\ d_i = u_i - u_i^C - \kappa_i(v_i' - v_i) \end{cases} \quad (9)$$

ここで、 (u_{0i}, v_{0i}) は $V_{R,i}^{rect}$ における画像中心、 β_i は $V_{R,i}^{rect}$ の焦点距離、 $\|\mathbf{t}_{cam,i}\|$ はステレオペア間の基線長、 v_i' は参照視点上の任意の垂直座標を表す。また、式 (9) は、図 3 のように、 $V_{R,i}^{rect}$ の X 軸正方向に C_i^{rect} があると仮定している。このとき、式 (9) を κ_i について解くことにより、

$$\kappa_i = -\frac{n_{Y,i} \|\mathbf{t}_{cam,i}\|}{\mathbf{n}_i \cdot \mathbf{M}_i} \quad (10)$$

を得る。 $V_{R,i}^{rect}$ と C_i^{rect} の位置関係が逆転すると、 κ_i の正負が逆転することより、式 (11) はより一般的に

$$\kappa_i = \frac{\mathbf{n}_i \cdot \mathbf{t}_{cam,i}}{\mathbf{n}_i \cdot \mathbf{M}_i} \quad (11)$$

と表せる。また、エピポーラ線が垂直座標軸と平行になるように平行化した場合においても、同様に考えると、スキューの傾きは式 (11) で表される。以上より、提案手法では、近傍視点におけるマッチングウィンドウを κ_i でスキューさせ、エピポーラ線と直行する座標軸のウィンドウ中心からのずれに応じて、ラインを切り出す中心座標を変化させることで、ステレオ画像間の局所的なスキューを軽減する。

3.3 画像変形の軽減を伴う POC による 3 次元復元

以上で述べた画像変形の軽減を伴う多視点ステレオのための POC に対して、プレーンスライピングを用いた奥行き探索と画像ピラミッドを用いた階層的探索を適用することで、多視点画像からの 3 次元復元を行う。このとき、物体表面の法線ベクトルは未知であるため、画像ピラミッド最上位層において法線ベクトルを変化させながらマッチン

グを繰り返し、最も POC 関数のピークが高くなるものを用いる。参照視点 V_R および近傍視点群 C の各画像とカメラパラメータを入力とし、 V_R における奥行きマップを生成する際の具体的な処理手順を以下に示す。

Step1: V_R および C から K 組の平行ステレオペア $V_{R,i}^{\text{rect}}, C_i^{\text{rect}}$ ($i = 0, \dots, K-1$) を生成する。また、各平行ステレオペアについて、画像を $1/2$ ずつ縮小した $l_{\max}$ 階層の画像ピラミッドを生成し、 l 階層における画像を $V_{R,i,l}^{\text{rect}}, C_{i,l}^{\text{rect}}$ とする。

Step2: V_R 上の画像座標 $\mathbf{m} = [u, v]$ について、各ステレオペアにおける視差の倍率 s_i を求める。

Step3: $l \leftarrow l_{\max}$ とし、 $l_{\max}$ 階層において、3 次元点 $\mathbf{M}$ の奥行きと法線ベクトル $\mathbf{n}$ を変化させながら多視点画像間の POC 関数 $\hat{r}_{\text{ave}}$ を計算し、最も $\hat{r}_{\text{ave}}$ のピーク値が高くなる $\mathbf{M}$ と $\mathbf{n}$ を求める。さらに、 $\hat{r}_{\text{ave}}$ のピーク座標から $\mathbf{M}$ を更新する。このとき、 $\mathbf{M}$ の奥行き変化の刻み幅は、ステレオ画像上でウィンドウサイズの $1/4$ に相当する奥行き変化量とする。また、本稿において、 $\mathbf{n}$ は、 V_R に正対した法線ベクトルを基準に、 $\pm\pi/8$ の範囲で X 軸および Y 軸回りに回転させた 9 つの法線ベクトルについて探索する。

Step4: $l \leftarrow l-1$ とし、最上位層で決定した法線ベクトル $\mathbf{n}$ を用いて、多視点画像間の POC 関数 $\hat{r}_{\text{ave}}$ を計算し、 $\hat{r}_{\text{ave}}$ のピーク座標から $\mathbf{M}$ を更新する。

Step5: $l \geq 1$ である限り、Step4 を繰り返す。 $l = 1$ の場合、3 次元点 $\mathbf{M}$ の奥行きを $\mathbf{m}$ における奥行きとする。

Step6: V_R 上の全ての座標について、 $\mathbf{m}$ を変化させながら、Step2 から Step5 の処理により 3 次元点の奥行きを計算し、奥行きマップを作成する。

以上の処理により求めた奥行きマップに対して、POC 関数のピーク値を用いた閾値処理による誤対応除去をすることで、高精度な 3 次元復元が可能である。さらに、参照視点 V_R を変えながら、異なる視点について 3 次元復元を行うことで、対象物体全体の 3 次元復元を行う。

4. 精度評価実験

本章では、Strecha らにより公開されているデータセット Fountain-P11 [12] を用いた実験により、提案手法の 3 次元復元精度を評価する。データセット Fountain-P11 には、多視点画像とそのカメラパラメータ、および復元対象の真値のメッシュモデルが含まれる。データセット Fountain-P11 に含まれる多視点画像を図 4 に示す。それぞれのデータセットのうち 1 枚を参照視点 V_R 、 V_R の近傍 2 枚を近傍視点 C とし、 V_R における奥行きマップの計算値と真値との比較を行う。各ピクセルにおける 3 次元点の奥行きの誤差として、次式で表される誤差率 e を用いる。

$$e = \frac{|Z_{\text{calculated}} - Z_{\text{ground truth}}|}{Z_{\text{ground truth}}} \quad (12)$$

ここで、 $Z_{\text{calculated}}$ は奥行きの計算値を、 $Z_{\text{ground truth}}$ は参照視点 V_R における真値のメッシュモデルの奥行きを表す。データセットに含まれる多視点画像は、 $3,072 \times 2,048$ ピクセルの画像であるが、その他に、 $1/2$ ($1,536 \times 1,024$ ピクセル) および $1/4$ (768×512 ピクセル) のサイズに縮小した画像を生成し、解像度の変化に対する比較も行う。

本稿では、従来手法として、NCC に基づく画像マッチングを用いたプレーンスイーピング (Goesele_2006 [4]) と、画像変形の軽減を行わない多視点ステレオのための POC (Sakai_2012 [10]) と比較する。Sakai_2012 についても、提案手法と同様に、プレーンスイーピングによる奥行き探索と画像ピラミッドを用いた階層的探索を組み合わせて奥行きマップを生成する。それぞれの手法について生成した奥行きマップについて、各ピクセルの誤差率を求めることで、3 次元復元精度を評価する。ここで、ある画像座標について、すべてのステレオペアで POC 関数のピーク値、または NCC の相関値が th_{corr} を下回った場合、奥行きが計算できないため、そのピクセルの 3 次元復元を行わない。また、画像マッチングの精度を評価するため、これ以外の誤対応除去は適用しない。従来手法と提案手法における各パラメータを以下のように設定する。

• Goesele_2006

Goesele_2006 では、奥行き変化の刻み幅 ΔZ が 3 次元復元精度に大きく影響する。そこで、本稿では、 ΔZ が最も基線長の長いステレオペア上で 1 または $1/10$ ピクセルとなるような、2 つの ΔZ について実験を行う。NCC のウィンドウサイズは 17×17 ピクセルとし、NCC の相関値を平均化する際の閾値 th_{corr} は 0.3 とする。

• Sakai_2012, 提案手法

POC 関数のピーク値の閾値 th_{corr} を 0.3、基準となるウィンドウの大きさ $w \times L$ を 32×17 ピクセルとする。ここで、1 次元 POC 関数の計算ではハニング窓をかけるため [13]、 32×17 ピクセルのウィンドウに含まれる情報量は、Goesele_2006 で用いる 17×17 ピクセルのウィンドウの情報量に相当する。また、画像ピラミッドの階層数は、オリジナルの画像サイズで 4 階層、 $1/2$ の画像サイズで 3 階層、 $1/4$ の画像サイズで 2 階層とする。

まず、ステレオ画像間の画像変形が大きくなるような、 V_R 、 C の組み合わせについて評価を行う。参照視点のカメラに対して対象物体が大きく傾いている場合として、10 番の画像を参照視点 V_R に、9, 11 番の画像を近傍視点 C とした場合の誤差率のヒストグラムを図 5 に示す。図 5 について、誤差率が 1 % 以上の点はヒストグラムの計算に用いているが、グラフには表示していない。このときの $1/2$

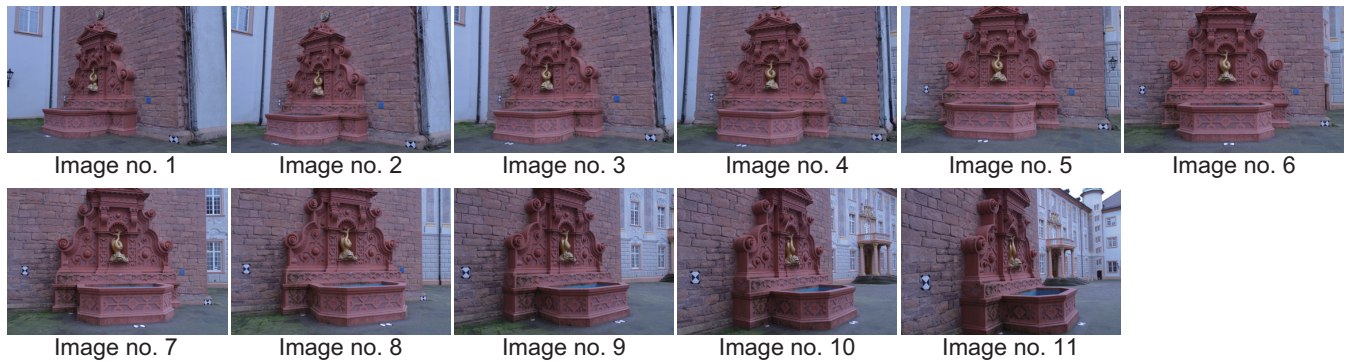


図 4 Fountain-P11 に含まれる多視点画像

サイズの画像における 3 次元復元結果を図 6 に示す。また、ステレオ画像間の基線長が長い場合として、9 番の画像を参照視点 V_R に、7, 11 番の画像を近傍視点 C とした場合の誤差率のヒストグラムを図 7 に、1/2 サイズの画像における 3 次元復元結果を図 8 に示す。図 5 および図 7 の誤差率ヒストグラムより、提案手法では、3 次元復元点のうち誤差率が小さい点が占める割合が多く、高精度な復元が行えている。このことは、図 6 および図 8 の 3 次元復元点群からも確認できる。Sakai_2012 のような画像変形の軽減を伴わない POC では、ステレオ画像間の変形が大きい場合、マッチングウィンドウ間の変形を平行移動のみに近似できないため、3 次元復元精度が低下する。これに対して、提案手法では、ステレオ画像間の局所的な変形を軽減することで、画像変形が大きい場合においても高精度な画像マッチングが可能であり、3 次元復元精度が向上している。また、従来の NCC を用いる手法に比べても高精度な 3 次元復元が行えている。

次に、参照視点 V_R を全視点で変化させながら生成した奥行きマップについて誤差を評価する。基線長が短くなるように、 h 番の参照視点に対して、 $h-1, h+1$ 番の画像を近傍視点とした場合の誤差率のヒストグラムを図 9 に、1/2 サイズの画像における 3 次元復元結果を図 10 に示す。また、基線長が長くなるように、 h 番の参照視点に対して、 $h-2, h+2$ 番の画像を近傍視点とした場合の誤差率のヒストグラムを図 11 に、1/2 サイズの画像における 3 次元復元結果を図 12 に示す。全視点の復元結果では、Sakai_2012 や提案手法のように、POC に基づく画像マッチングを用いた手法が高精度な 3 次元復元が行えている。さらに、提案手法では、画像間の変形が大きい場合における精度が改善しているため、Sakai_2012 と比べても高精度な 3 次元復元が行えている。

5. まとめ

本稿では、ステレオ画像間の局所的な画像変形をマッチングウィンドウの拡大縮小とスキューにより軽減することで、多視点ステレオのための POC に基づく画像マッチン

グの高精度化を行なった。提案手法に対して、プレーンスライピングによる奥行き探索と画像ピラミッドを用いた階層的探索を組み合わせることで、ステレオ画像間の変形が大きい場合においても、高精度な 3 次元復元を実現した。今後は、提案手法を用いた、より効率的な多視点ステレオアルゴリズムの開発を行う。

参考文献

- [1] Szeliski, R.: *Computer Vision: Algorithms and Applications*, Springer-Verlag New York Inc. (2010).
- [2] Seitz, S. M., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R.: A comparison and evaluation of multi-views stereo reconstruction algorithms, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 519–528 (2006).
- [3] Strecha, C., von Hansen, W., Gool, L. V., Fua, P. and Thoennessen, U.: On benchmarking camera calibration and multi-view stereo for high resolution imagery, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 1–8 (2008).
- [4] Goesele, M., Curless, B. and Seitz, S. M.: Multi-view stereo revisited, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 2402–2409 (2006).
- [5] Goesele, M., Snavely, N., Curless, B., Hoppe, H. and Seitz, S. M.: Multi-view stereo for community photo collections, *Proc. Int'l Conf. Computer Vision*, pp. 1–8 (2007).
- [6] Bradley, D., Boubekeur, T. and Heidrich, W.: Accurate multi-view reconstruction using robust binocular stereo and surface meshing, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 1–8 (2008).
- [7] Campbell, N. D. F., Vogiatzis, G., Hernandez, C. and Cipolla, R.: Using multiple hypotheses to improve depth-maps for multi-view stereo, *Proc. European Conf. Computer Vision*, pp. 766–779 (2008).
- [8] Hiep, V. H., Keriven, R., Labatut, P. and Pons, J. P.: Towards high-resolution large-scale multi-view stereo, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 1430–1437 (2009).
- [9] Furukawa, Y. and Ponce, J.: Accurate, dense, and robust multiview stereopsis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 32, No. 8, pp. 1362–1376 (2010).
- [10] Sakai, S., Ito, K., Aoki, T., Masuda, T. and Unten, H.: An Efficient Image Matching Method for Multi-View Stereo, *Proc. Asian Conf. Computer Vision*, pp. 1–8

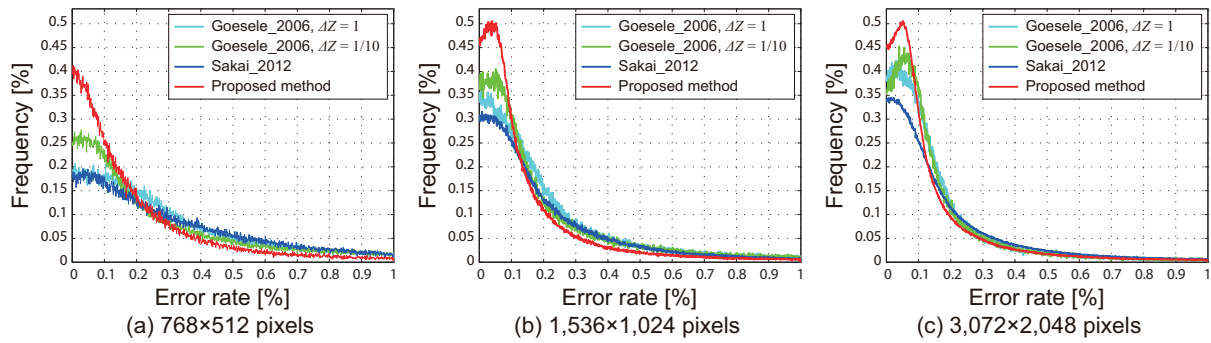


図 5 カメラに対して対象物体が大きく傾いている場合の誤差率ヒストグラム



図 6 カメラに対して対象物体が大きく傾いている場合の 3 次元復元結果: (a) Goesele_2006 $\Delta Z = 1/10$ ピクセル, (b) Sakai_2012, (c) 提案手法, (d) 真値のメッシュモデル

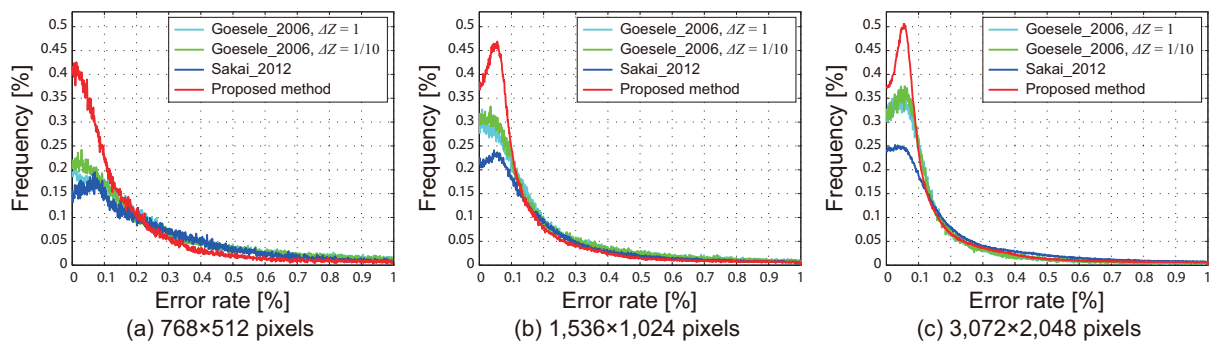


図 7 ステレオ画像間の基線長が長い場合の誤差率ヒストグラム



図 8 ステレオ画像間の基線長が長い場合の 3 次元復元結果: (a) Goesele_2006 $\Delta Z = 1/10$ ピクセル, (b) Sakai_2012, (c) 提案手法, (d) 真値のメッシュモデル

- (2012).
- [11] Takita, K., Aoki, T., Sasaki, Y., Higuchi, T. and Kobayashi, K.: High-accuracy subpixel image registration based on phase-only correlation, *IEICE Trans. Fundamentals*, Vol. E86-A, No. 8, pp. 1925–1934 (2003).
 - [12] Strecha, C.: Multi-View Evaluation, <http://cvlab.epfl.ch/data/>.
 - [13] 柴原琢磨, 沼 徳仁, 長嶋 聖, 青木孝文, 中島 寛, 小

- 林孝次: 1 次元位相限定相関法に基づくステレオ画像の高精度サブピクセル対応付け手法, 電子情報通信学会論文誌 D, Vol. J91-D, No. 9, pp. 2343–2356 (2008).
- [14] Sakai, S., Ito, K., Aoki, T. and Unten, H.: Accurate and dense wide-baseline stereo matching using SW-POC, *Proc. Asian Conf. Pattern Recognition*, pp. 335–339 (2011).

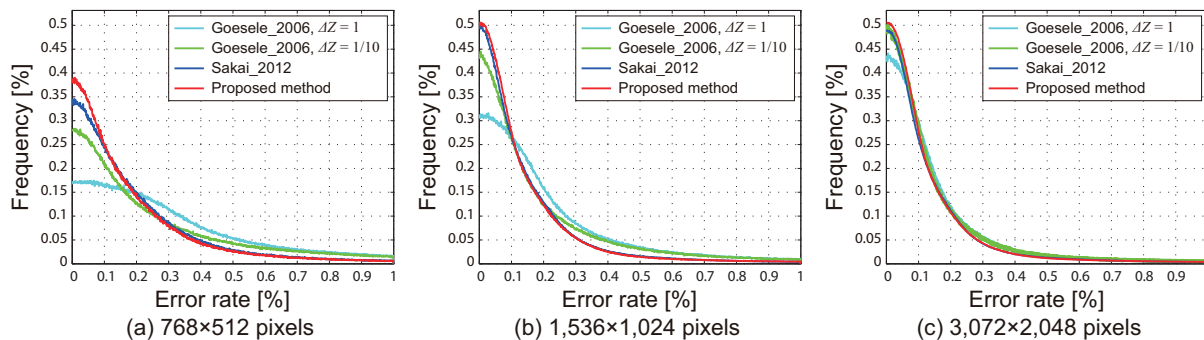


図 9 基線長の短いステレオペアを選択した場合の誤差率ヒストグラム



図 10 基線長の短いステレオペアを選択した場合の 3 次元復元結果: (a) Goesele_2006 $\Delta Z = 1/10$ ピクセル, (b) Sakai_2012, (c) 提案手法, (d) 真値のメッシュモデル

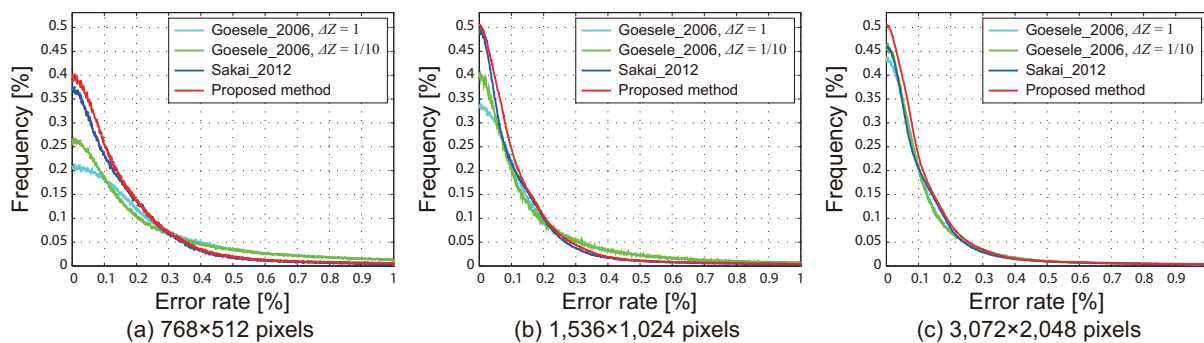


図 11 基線長の長いステレオペアを選択した場合の誤差率ヒストグラム



図 12 基線長の長いステレオペアを選択した場合の 3 次元復元結果: (a) Goesele_2006 $\Delta Z = 1/10$ ピクセル, (b) Sakai_2012, (c) 提案手法, (d) 真値のメッシュモデル