

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手法と その性能評価

酒井 修二† 伊藤 康一† 青木 孝文† 増田 智仁‡ 運天 弘樹‡

東北大学 大学院情報科学研究科† 凸版印刷株式会社‡

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]。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次元点の推定、(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、3次元点を推定するための画像マッチング手法は、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。

従来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[2]。これらのアルゴリズムでは、3次元点の座標を離散的に変化させながら繰り返し画像マッチングを行い、最もマッチングスコアの高くなる3次元座標を復元点とする。このとき、高精度な3次元復元を行うためには、非常に細かい刻み幅で座標を変化させる必要があり、計算コストが膨大になるという問題がある。

これに対して、本稿では、多視点ステレオのための位相限定相関法 (Phase-Only Correlation: POC) [3] に基づく画像マッチング手法を提案する。提案手法では、1回の画像マッチングにより得られるPOC関数から直接3次元座標を推定するため、従来手法と比べて非常に少ない回数の画像マッチングで3次元復元を行うことが可能である。一般に公開されている評価用データセット[4]を用いた実験を通して、従来のNCCに基づく画像マッチング手法に比べ、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す。

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

POCは、画像の位相情報のみを用いた画像マッチング手法である[3]。2つの画像をフーリエ変換し、それ

ぞれの位相スペクトルからPOC関数を計算する。2つの画像間の変形が平行移動のみの場合、POC関数の理論的なピークモデルを定義することが可能である[3]。このとき、POC関数のピークの高さは画像間の類似度に、POC関数のピーク座標は画像間の平行移動量に相当する。POC関数の計算値に理論的なピークモデルを当てはめ、ピーク座標を実数値で求めることで、2つの画像間の平行移動量をサブピクセル精度で推定する。

POCを多視点ステレオに適用する場合、1枚の参照視点と k 枚の近傍視点から k 組のステレオペアを作り、それぞれのステレオ画像間の局所的なウィンドウマッチングにPOCを用いる[5]。多視点ステレオにおいて、参照視点における視線上の3次元点の移動は、近傍視点においてエピポーラ線上の平行移動となる。この平行移動量をPOC関数のピーク座標から計算することで、3次元点の移動量を推定できる。つまり、ある3次元候補点について、ステレオ画像上の投影点周辺に設定されたウィンドウ間でPOC関数を計算することで、3次元候補点が真の3次元点からどれくらい離れているかを1回の画像マッチングで推定することが可能である。このとき、提案手法では、 k 組のステレオペアそれぞれで計算されるPOC関数を正規化視差の概念に基づき統合することで、3次元座標推定の精度とロバスト性を向上させる[5]。さらに、提案手法をブレンスレーピングや階層的探索と組み合わせ、非常に少ないマッチング回数で高精度な3次元復元を行う。

3 性能評価実験

3次元復元精度と計算コストの観点から、提案手法の性能評価を行う。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多視点ステレオの1つであるGoeseleらの手法[2]を用いる。Goeseleらの手法では、3次元点の座標を変化させる刻み幅 ΔZ が精度と計算コストに大きく影響するため、 ΔZ が最も基線長の長いステレオ画像上で1, 1/2, 1/5, 1/10 pixelとなる4つの場合について実験を行う。

まず、公開データセットFountain-P11, Herz-Jesu-P25[4]を用いて3次元復元精度を評価する。データセットに含まれる多視点画像のすべての視点について、各画像

POC-Based Image Matching for Multi-View Stereo and Its Performance Evaluation

†Shuji SAKAI †Koichi ITO †Takafumi AOKI

‡Tomohito MASUDA ‡Hiroki UNTEN

†Graduate School of Information Sciences, Tohoku University

‡Toppan Printing Co., Ltd.

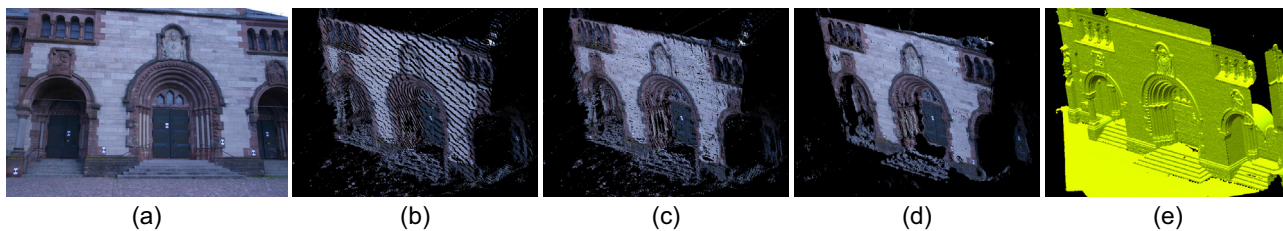


図 1: 3次元復元結果の例: (a) 参照視点画像, (b) Goesele らの手法 $\Delta Z = 1$ pixel による結果, (c) Goesele らの手法 $\Delta Z = 1/10$ pixel による結果, (d) 提案手法による結果, (e) 真値のメッシュモデル

表 2: 1 点の奥行き探索に要する計算コスト

	Additions	Multiplications	Divisions	Square roots
Goesele, $\Delta Z = 1$ pixel	75,140	31,246	578	578
Goesele, $\Delta Z = 1/2$ pixel	150,280	62,492	1,156	1,156
Goesele, $\Delta Z = 1/5$ pixel	357,700	156,230	2,890	2,890
Goesele, $\Delta Z = 1/10$ pixel	751,400	312,460	5,780	5,780
Proposed method	40,000	34,496	2,176	1,088

表 1: 平均誤差率 [%]

	Fountain	Herz-Jesu
Goesele, $\Delta Z = 1$ pixel	0.2924	0.2751
Goesele, $\Delta Z = 1/2$ pixel	0.2613	0.2496
Goesele, $\Delta Z = 1/5$ pixel	0.2509	0.2429
Goesele, $\Delta Z = 1/10$ pixel	0.2493	0.2422
Proposed method	0.1864	0.1855

座標における 3 次元点の奥行きを計算し、真値の奥行きと比較する。このとき、参照視点 1 枚につき、Fountain-P11 では近傍視点を 2 枚、Herz-Jesu-P25 では近傍視点を 4 枚用いる。また、それぞれの画像は、オリジナルの $1/4$ サイズ (768×512 pixel) に縮小する。本稿では、奥行き計算値 $Z_{\text{calculated}}$ と奥行き真値 $Z_{\text{ground truth}}$ との誤差として、次式で定義される誤差率 e を用いる。

$$e = \frac{|Z_{\text{calculated}} - Z_{\text{ground truth}}|}{Z_{\text{ground truth}}} \times 100 \text{ [%]} \quad (1)$$

図 1 に、Herz-Jesu-P25 の視点番号 19 の画像を参照視点として 3 次元復元した場合の結果を示す。Goesele らの手法では、 ΔZ が小さい場合に、復元結果が階段状になっていることが目視で確認できる。表 1 にそれぞれのデータセットに対する誤差率の平均値を示す。平均誤差率の計算では、誤差率 1% 以上の点を誤対応点とし、復元結果から誤対応点を除去して平均値を計算する。表 1 より、Goesele らの手法では ΔZ が小さいほど平均誤差率が小さくなっている。一方、提案手法では、Goesele らの手法 $\Delta Z = 1/10$ pixel と比べても平均誤差

率が小さく、より高精度な 3 次元復元が行えている。

次に、計算コストを評価する。本稿では、参照視点上のある 1 点の奥行きを計算するのに必要な加算、乗算、除算、平方根の計算の演算回数を計算コストとする。表 2 にそれぞれの手法における演算回数を示す。表 2 より、提案手法の計算コストは、Goesele らの手法 $\Delta Z = 1$ pixel から $\Delta Z = 1/2$ pixel 程度であり、Goesele らの手法 $\Delta Z = 1/10$ pixel と比べると非常に少ない。

4 まとめ

本稿では、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手法を提案し、その性能評価を行なった。提案手法を用いることで、従来の NCC に基づく画像マッチング手法に比べ、少ない計算コストで高精度な 3 次元復元が可能であることを示した。

参考文献

- [1] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag New York Inc., 2010.
- [2] M. Goesele, B. Curless, and S. M. Seitz. Multi-view stereo revisited. *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 2402–2409, June 2006.
- [3] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi. High-accuracy subpixel image registration based on phase-only correlation. *IEICE Trans. Fundamentals*, Vol. E86-A, No. 8, pp. 1925–1934, August 2003.
- [4] C. Strecha. Multi-view evaluation. <http://cvlab.epfl.ch/data/>.
- [5] S. Sakai, K. Ito, T. Aoki, T. Masuda, and H. Unten. An efficient image matching method for multi-view stereo. *Proc. Asian Conf. Computer Vision*, pp. 1–8, November 2012.