# 正距円筒オプティカルフローパターンを均等化した E-CNN による 全天球カメラの回転推定の精度向上

○ Kim Dabae, Pathak Sarthak, Moro Alessandro, 小松 廉, 山下 淳, 淺間 一 (東京大学)

# E-CNN : Uniformization of Equirectangular Optical Flow Patterns for Accuracy Improvement of Spherical Camera Rotation Estimation

○ Dabae KIM, Sarthak PATHAK, Alessandro MORO, Ren KOMATSU, Atsushi YAMASHITA and Hajime ASAMA (The University of Tokyo)

Abstract : Spherical cameras acquire all-round information, making motion estimation effective. Convolutional Neural Network (CNN) is robust for such tasks but cannot be applied to spherical images due to their non-planarity. When spherical images are projected as planar equirectangular images, the unequal distortion results in low accuracy. In this paper, we propose a novel, distortion-resistant E-CNN, which robustly estimates spherical camera rotation.

## 1. 序論

未知環境の探索において,移動ロボットの運動推定は重 要な課題である.運動推定の方法として,移動ロボットに カメラを装着し,周囲環境の変動から推定を行うことは多 く用いられる.一般的なカメラを用いる場合,カメラ視野 の狭さの影響で,障害物などによっては正確な運動推定が できない可能性がある.一方,360°の視野角を持つ全天球 カメラは,一般的なカメラより広い角度の情報が取得でき, 運動推定の精度向上が期待できる.また,カメラ視野の一 部が妨げられても,周囲環境から運動推定を行い続けるこ とができる.つまり,全天球カメラは運動推定において有 効なデバイスであると言える.

カメラの運動推定を行う一般的な手法として,取得し た画像から特徴量を設計するフィーチャーベースト手法 が挙げられる<sup>1)2)</sup>.しかし,この手法では特徴量の設計 が必要となる.また,環境によっては運動推定が失敗す る場合がある.近年,カメラの運動推定を行う手法とし て,ラーニングベースト手法が提案されている.この手 法では,大量のデータを学習させることで,環境変動にロ バストな運動推定が可能であることが示されている.その 中でも,画像に対しての畳み込みを行って特徴を抽出する CNN (Convolutional Neural Network)は,カメラの運 動推定において,環境変動に対するロバスト性を確認して いる<sup>3)4)5)</sup>.

全天球カメラは球面上に投影された像を画像として取 得するため、平面画像に対する2次元の畳み込みを行う CNN をそのまま用いることは困難である.そこで、全天 球カメラは球面画像の以外に、地球儀に対する世界地図の ように、球面上に投影された像を平面上に引き伸ばした正



Fig. 1: Spherical and Equirectangular Images

距円筒画像に変換することができる(Fig. 1).しかし,像 が引き伸ばされた正距円筒画像には,Fig. 1の上端と下端 のような情報の歪みが生じる.CNNを用いて全天球カメ ラの運動推定を行うとき,正距円筒画像をそのまま畳み込 むと,情報の歪みによって推定精度が低下する問題が生じ る.つまり,この歪みの問題を解決することは必須不可欠 である.

正距円筒画像を用いるときに生じる歪みを考慮すること は、全天球カメラの運動推定における精度向上に繋がる. そこで、本研究では、正距円筒画像の歪みを考慮した E-CNN (Equirectangular-Convolutional Neural Network) を提案し、全天球カメラの3自由度の回転推定を行うこと を目的とする.具体的には、正距円筒画像の不均等な歪み の影響を受ける回転のパターンを均等化させることで、全 天球カメラの回転推定における精度向上を目指す.回転の パターンには、カメラの2フレーム間の画像から求めたデ ンスなオプティカルフローを用いる.これについては 2.1 節で詳しく説明する.



Fig. 2: Dense Optical Flow



Fig. 3: Optical Flow Patterns: Lined and Curved

# 2. 提案手法

## 2.1 デンスなオプティカルフロー

本研究では、全天球カメラから取得した2フレーム間の 正距円筒画像からデンスなオプティカルフロー(Fig. 2) を生成する.ここで、デンスなオプティカルフローとは、 前後のフレームにおいて、すべての画素の移動をベクトル として表したものである.

本研究において、2 フレーム間のデンスなオプティカル フローを入力とすることで、回転による全画素の変化が学 習できる.また、デンスなオプティカルフローを使用する ことで、物体のテクスチャ情報に依存しない、異なる環境 に対するロバスト性が期待できる.

オプティカルフローは垂直成分と水平成分で表現するこ とができ、本研究ではそれぞれを分けてデータセットの作 成を行う. これについては 3.1 節で詳しく説明する.

### 2.2 オプティカルフローパターンの均等化

2.1 節で述べたように、デンスなオプティカルフローは 画素の移動をベクトルで表したものである.ただし、正



Standard axis: Pitch



Standard axis: Roll Fig. 4: Uniformized Optical Flow Patterns

距円筒画像の歪みはオプティカルフローに影響する.例 えば、全天球カメラの基準軸が Fig. 3 のようであるとき、 ヨー方向周りの回転によるオプティカルフローは直線のパ ターンを示す(Fig. 3).一方、ロールとピッチ方向周り の回転によるオプティカルフローは曲線のパターンを示す (Fig. 3).つまり、ロール・ピッチ・ヨーのすべての方向に 対して回転する場合、オプティカルフローは直線のパター ンが1つ、曲線のパターンが2つ混ざっていることになる. 例えば、ロール方向を基準軸とする場合、ロール方向周り の回転によるオプティカルフローは直線のパターンであり、 ピッチとヨー方向周りの回転によるオプティカルフローは 曲線のパターンである.また、ピッチとヨー方向を基準軸 とする場合にも同様なことが言える.

1章で述べたように,全天球カメラの正距円筒画像には 歪みの問題があり,そのまま用いても高精度な推定結果が 得られない.本研究では,この問題を解決するために,各 基準軸に対する回転のパターンを均等化することを考える. 具体的には,Table 1の状態のように,各基準軸に対する 回転軸の直線・曲線のオプティカルフローパターンを均等 にする.

そのために, Fig. 3の基準軸において, ロールとピッチ



Fig. 5: E-CNN

Table 1: Optical Flow Patterns Uniformization

回転軸 基準軸	ロール	ピッチ	-Е
ロール	直線	曲線	曲線
ピッチ	曲線	直線	曲線
-E	曲線	曲線	直線

をヨー方向に合わせて軸を 90° 回転させることを考える. 軸を 90° 回転させることでその軸は基準軸となる.具体 的には,2フレームをロール方向に 90° 回転させ,回転し たフレーム間のデンスなオプティカルフローを算出する. ピッチ方向にも同様に行う.結果的に,3つのデンスなオ プティカルフローが生成される (Fig.4).

以上のように、オプティカルフローパターンが均等化 された入力データを用いて CNN による特徴抽出を行 うネットワークを本研究では E-CNN (Equirectangular-Convolutional Neural Network) と呼ぶ.

### 2.3 E-CNN

提案手法の E-CNN は,2 フレーム間の正距円筒画像に おけるオプティカルフローパターンの均等化(2.2節)を 行うネットワークとオプティカルフローの特徴抽出を行う ネットワークによって構成されている.

本研究では,正距円筒画像から算出したオプティカル フローの特徴抽出を行うネットワークとして CNN を用い る.用いる CNN ネットワークは, Fig. 5 に示すように, Conv1<sub>[3×3×128]</sub>, Pool1<sub>[2×2]</sub>, Conv2<sub>[3×3×128]</sub>, Pool2<sub>[2×2]</sub>, Conv3<sub>[3×3×256]</sub>, Pool3<sub>[2×2]</sub>, fc<sub>[256]</sub>, fc<sub>[4]</sub> で構成されて いる.

本研究では、回転推定を回帰問題として解く.そのため に、全結合層の回帰器では真値と推定値とのユークリッド 距離を損失関数として、学習におけるパラメータの更新を 行う.損失関数は以下の式(1)で表される.

$$\mathcal{L}(\boldsymbol{I}) = \|\Delta \hat{\boldsymbol{q}} - \Delta \boldsymbol{q}\|_2. \tag{1}$$

式 (1) は、2 フレーム間のデンスなオプティカルフロー パターンの均等化を行った入力 **I** において、ネットワーク が推定したクォータニオンで表す回転量  $\Delta q = (\Delta w, \Delta a, \Delta b, \Delta c)$  と、回転量の真値  $\hat{q}$  の差分の二乗誤差を表してい る. つまり、この損失関数が最小になるようにパラメータ を更新する回帰を行う.

## 3. 実験

### 3.1 データセットの作成

データセットは3次元コンピュータグラフィックスソフ トウェアである Blender を使用して作成した. Blender 環 境とカメラ視点を Fig. 6 に示す.

まず, Blender 上の屋外の都心<sup>7)</sup> における学習用の 10 シーン(1~10)を設定し, 1~10 シーン以外の検証・テ スト用の 10 シーン(11~20)を別途に設定した. 学習用 と検証・テスト用のシーンに別々にする理由は, 学習した シーンに対する過学習を防止し, 学習に一般性を保たせる ためである.



Fig. 6: Blender Environment (left) and Camera Viewpoint (right)

次に,全天球カメラ画像を大きさ 100×200 の正距円筒 画像として取得した.学習データセットについては,学習 に一般性を持たせるために,取得するシーン(1~10)と 初期姿勢をランダムに設定した.また,初期姿勢からの回 転量は各回転軸に対して 0°~10°に制限し,0.5°刻みとし て計 9,261 枚のオプティカルフローを学習データとした. 回転量を制限する理由は,オプティカルフローが算出でき る角度と,各フレームは移動量は微小であることを考慮し たためである.

検証・テストデータについては、学習データセットと同様にシーン(11~20)と初期姿勢をランダムに設定した. また、回転量を各回転軸に対して 0°~10°に制限し、ラン ダムに設定して計 1,000 枚のオプティカルフローを各々用 意した.以上で説明した各データセットの構成については Table 2 に示す.

Table 2: Datasets Composition

種類項目	学習データ	検証データ	テストデータ
シーン番号	$1 \sim 10$	$11 \sim 20$	$11 \sim 20$
枚数 [枚]	9,261	1,000	1,000

デンスなオプティカルフローの算出には, DeepFlow<sup>6)</sup> を用いる.この手法では,移動前後の画素の輝度値は不変 であり,全体的なオプティカルフローは滑らかであること を仮定し,特徴点のマッチングを行うことでオプティカル フローを推定している.

最後に、2.2 節で述べたオプティカルフローのパターン を均等化を行うために、ロールとピッチ方向に各々90°回 転させ、2 フレーム間のオプティカルフローを生成した. 以上で生成した3つのオプティカルフローを垂直成分と水 平成分に分けて6 チャンネルの配列として保存した.

#### 3.2 推定実験

2.3 節で提案した E-CNN ネットワークを用いて回転推 定の実験を行った.活性化関数には ReLU<sup>8)</sup>を用い,パ ラメータの更新方法には Adam オプティマイザー<sup>9)</sup>を採 用した.また,バッチサイズは 128,学習率は 0.001,エ



ポックは 100 で学習を行った. 学習・検証・テストにおい て、CPU は NVIDIA 制の CoForge CTY 1080Tiv 2 を

て, GPU は NVIDIA 製の GeForce GTX 1080Ti×3 を, CPU は Intel 製の Xeon E5-1650 v4 を使用した.

推定実験では,提案手法であるオプティカルフローパ ターンの均等化による精度向上を確認するために,均等化 前後の2種類をデータセットを用いて行った.なお,デー タセットや活性化関数,パラメータの更新方法,ハイパー パラメータなどは同条件で行った.

推定結果を Table 3 に示す. この表では, 1,000 枚のテ ストデータにおける推定誤差を角度として表したときの中 央誤差と平均誤差を示している.

Table 3: Estimation Results

種類 誤差	均等化前	均等化後(提案手法)
中央誤差 [°]	0.270	0.198
平均誤差 [°]	$0.278 \pm 0.106$	$0.211 \pm 0.100$

また,1,000 枚のテストデータに対する均等化前後の推 定誤差のヒストグラムを Fig. 7 に示す.縦軸はテストデー タのサンプル数,横軸は推定誤差を示している.

実験では、高精度な回転推定の結果が得られており、さらに、提案手法による約 24.1%の精度向上が確認できた. また、学習データと異なるシーンに対するテストデータを 用いることで、学習の汎化能力を確認した.

### 4. 結論

本研究では、全天球カメラから取得した正距円筒画像の 歪みによって生じる不均等なオプティカルフローパターン を均等化する E-CNN を構築して全天球カメラの 3 自由 度の回転推定を行った.推定実験では、提案手法による精 度向上を確認した.今後の課題としては、実際のカメラに よって取得したデータに対する検証と全天球カメラの 6 自 由度の運動(回転・並進)推定が挙げられる.

# 謝辞

本研究の一部は、内閣府総合科学技術・イノベーション 会議の戦略的イノベーション創造プログラム(SIP)「イン フラ維持管理・更新・マネジメント技術」(管理法人:国 立研究開発法人新エネルギー・産業技術総合開発機構) に よって実施された.

# 参考文献

- 1) G. Klein and D. Murray: "Parallel Tracking and Mapping for Small AR Workspaces", Proceedings of the 6th IEEE/ACM International Symposium on
- Mixed and Augmented Reality, pp. 225–234, 2007. 2) R. Mur-Artal, J. Montiel and J. Tardos: "ORB-SLAM: A Versatile and Accurate Monocular SLAM System", *IEEE Transactions on Robotics*, Vol. 31, No. 5, pp. 1147–1163, 2015.
- 3) R. Clark, S. Wang, H. Wen, A. Markham and N. Trigoni: "VINet: Visual-Inertial Odometry as a Sequence-to-Sequence Learning Problem", Proceedings of the 31th AAAI Conference on Artificial In*telligence*, pp. 3995–4001, 2017. 4) G. Constante, M. Mancini, P. Valigi and T. Cia-
- rfuglia: "Exploring Representation Learning With

Vol. 1, No. 1, pp. 18–25, 2016.

- 5) M. Qiao and Z. Wang: "Learning the Frame-2-Frame Ego-Motion for Visual Odometry with Convolutional Neural Network", Proceedings of the 2th Chinese Conference on Computer Vision, pp. 500-511, 2017.
- 6) P. Weinzaepfel, J. Revaud, Z. Harchaoui and C. Schmid: "DeepFlow: Large Displacement Optical Flow with Deep Matching", Proceedings of the 2013 IEEE International Conference on Computer Vision, pp. 1385–1392, 2013.
- 7) Z. Zhang, H. Rebecq, C. Forster and D. Scara-muzza: "Benefit of Large Field-of-View Cameras for Visual Odometry", Proceedings of the 2016 IEEE International Conference on Robotics and Automation, pp. 801-808, 2016.
- 8) V. Nair and G. Hinton: "Rectified Linear Units Improve Restricted Boltzmann Machines", Proceedings of the 27th International Conference of Machine Learning, pp. 807-814, 2010.
- 9) D.P. Kingma and J. Ba: "Adam: A Method for Stochastic Optimization", Proceedings of the 3rd International Conference for Learning Representations, pp. 1–13, 2015.