

ニューラルネットワークを用いた3次元動き解釈システム のカメラ配置に関する検討*

4F-1

矢沢 伸行[†] 全 炳東[‡] 宮内 新[†] 宮内 ミナミ[§]
武蔵工業大学[†] 東京商船大学[‡] 産能大学[§]

1 はじめに

これまで物体の運動を画像から得るために多くの手法が提案されてきている。その一つに画像からオプティカルフロー (OF:Optical Flow) を求め、OFを解釈することで物体の運動を得る手法がある。OFを用いる手法のうち、複素バックプロパゲーションアルゴリズム^[1]を応用したニューラルネットワークを用いてOFを解釈する手法^[2]は他の手法に比べノイズに強く、実用的である。我々は同手法を応用した三眼視による3次元動き解釈手法^[3]を提案し、3次元での動きが高い精度で解釈できることを確認している。

しかし、3次元運動を解釈する際に用いた3次元動き解釈システムでは、対象物体を撮影するためのカメラを互いに垂直な3軸方向に配置する必要があった。この制約は提案された動き解釈手法の応用可能な分野を著しく制限している。本研究ではニューラルネットワークの学習方法を変更することで、カメラ配置の自由度を高めることが出来ることを示す。

2 3次元動き解釈システム

3次元動き解釈システムの構成を図1に示す。カメラで撮影した画像から抽出した任意の大きさのOFは、正規化ネットワークによってN×Nの大きさに正規化する。2次元動き解釈ネットワークは入力されたN×NのOFからX、Y方向の並行移動と回転、拡大縮小の4つの2次元動き

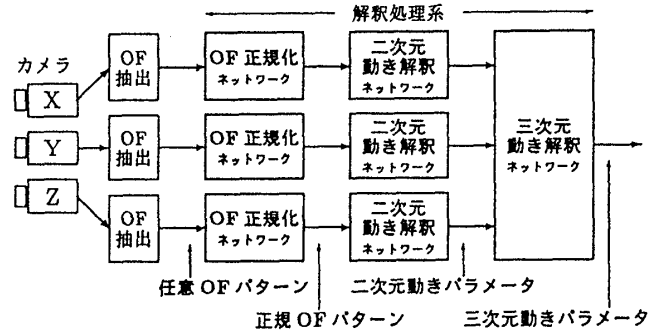


図 1: 3次元動き解釈システム

パラメータを出力する。3台のカメラから得られた合計12の2次元動きパラメータは3次元動き解釈ネットワークに入力される。3次元動き解釈ネットワークは対象物体の3次元上でのX、Y、Z方向への並行移動と回転の6つの3次元動きパラメータを出力する。この3次元動き解釈システムは任意の3次元運動をしている対象物体に対し、ノイズのない画像であれば平均誤差2.5%、200%のノイズを加えた画像に対しても平均誤差22%の精度で3次元運動を解釈出来ることが確かめられている。

しかしこの3次元動き解釈システムは、対象物体を撮影するためのカメラが互いに垂直な3軸方向に配置されていなければならないという制約がある。この制約は3次元動き解釈システムを応用することのできる場面を著しく制限しており、改善する必要がある。

3 改善方法

図2に従来のカメラと物体座標系との関係を示す。これまでは図2に示す物体座標系の原点付近で運動している物体を互いに垂直な3軸方向に配置されたカメラで撮影していた。ここでは3台のカメラのうちカメラZについて、カメラZとカメラX、カメラYとの間の距離を等しく保

*Camera positions for 3D motions interpret system using Neural Networks'

[†]Nobuyuki Yazawa, Arata Miyauchi
Musashi Institute of Technology
[‡]heitou zen
Tokyo University of Mercantile Marine
[§]Minami Miyauchi
Sanno College

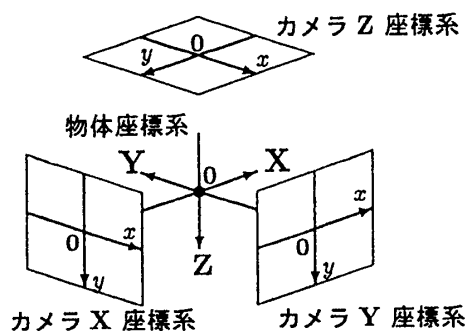


図 2: 従来のカメラと物体座標系との関係

しながらカメラ Z をカメラ X、カメラ Y へ近付けていくことを考える。実験ではカメラ Z の光軸とカメラ X、カメラ Y の作る平面との角度を 0° から 90° の範囲で 15° ずつずらしていった場合について、シミュレーションを行った。

ニューラルネットワークの学習データを変更することでカメラ配置の自由度を増すには 2 つの方法が考えられる。2 次元動き解釈ネットワークに任意の方向から見た時の OF を学習させ、3 次元動き解釈ネットワークの学習データは変更せずに解釈させる方法と、2 次元動き解釈ネットワークの学習データは換えずに任意の方向から見た時の OF を入力し、その出力結果を 3 次元動き解釈ネットワークに学習させる方法である。本研究では両方の方法で実験を行った。

4 実験結果

4.1 2次元動き解釈ネットワークの学習データを変更した場合

2 次元動き解釈ネットワークに対し、カメラ Z から得た OF を入力すると、カメラ Z が物体座標系の Z 軸上に配置されていた場合に得られる 2 次元動きパラメータを出力するように学習を行った。しかし、カメラ Z が Z 軸上に配置されている場合を除いて 2 次元動き解釈ネットワークは収束せず、発散してしまった。

4.2 3次元動き解釈ネットワークの学習データを変更した場合

3 次元動き解釈ネットワークに対し、カメラ配置を変更した時の OF を 2 次元動き解釈ネットワークが解釈した結果を 3 次元動き解釈ネット

ワークに学習させた。学習は、教師データと 3 次元動き解釈ネットワークの解釈結果との 2 乗平均誤差 (MSE) が 0.001 になるまで行った。

学習終了後、3 次元動き解釈システムに対して学習していない動きを示す OF を 25 パターン入力し、どのような精度で解釈できるかを調べた。ただし、OF にノイズは含まれていない。その時の 3 次元動き解釈システムの解釈結果を表 1 に示す。表 1 より、カメラ Z の光軸とカメラ X、カメラ Y の作る平面との角度が浅くなるのに従い平均解釈誤差が増えていることがわかる。しかし、最大でも平均解釈誤差 8.3% で解釈できた。

表 1: 3次元動き解釈ネットワークの解釈結果

角度 [$^\circ$]	90	75	60	45	30	15	0
平均誤差 [%]	2.7	4.0	4.7	5.6	7.7	7.0	8.3

5 まとめ

3 次元運動解釈システム内のニューラルネットワークの学習データを換えることでカメラ配置の自由度を高めるためることが出来るかどうかを検討した。

2 次元解釈ネットワークの学習データを変更する方法では学習が収束しなかったが、3 次元解釈ネットワークの学習データを変更する方法では、カメラ配置を変更してもノイズを含まない画像に対し平均解釈誤差 8.3% 以内の精度で解釈できることを確認した。今後の課題は、ノイズを含んだ画像に対する解釈精度を確認することと、さらにもう一台のカメラの配置も変更し、カメラ配置の自由度をより高めることである。

参考文献

- [1] 新田, 古谷: “複素バックプロパゲーション学習” 情報処理学会論文誌 Vol.32, No.10, pp.1319-1329, 1991
- [2] M. Miyauchi, M. Seki, A. Watanabe, A. Miyauchi: “Interpretation of Optical Flow through Complex Neural Network” Proc. of International Workshop on Artificial Neural Networks (IWANN'93-Barcelona), pp.645-650, 1993
- [3] 渡辺, 宮内, 宮内: “ニューラルネットワークを用いた物体の三次元動き解釈手法の研究” 第 16 回情報理論とその応用シンポジウム (SITA'93) 予稿集 (2), pp.529-532, 1993