

ニューラルネットワークによるオプティカルフローの
三次元動き解釈のための一提案*

4C-8

渡部 朗†,

宮内 新†,

宮内 ミナミ††

武蔵工業大学†,

産能大学††

1 はじめに

我々はこれまでに、オプティカルフロー(OF—画像から計算した動きの場合)の解釈に複素バックプロパゲーション(複素BP)アルゴリズム[2]を応用したニューラルネットワークによるOFの解釈法について検討を行い、良い精度の結果が得られる事を確認している[4].

本稿では、同手法を拡張してステレオ画像法の原理と組み合わせ、二眼視または三眼視で撮影した同一物体についての画像より別々に求めた、複数のOFの解釈結果を元に、対象物体の三次元動きを解釈するための手法について考察する.

一般的な手法として考えられる、複数のOFの解釈結果の連立解により三次元動きを計算する方法は非効率的であることから、これに代わって三次元ベクトルバージョンBPアルゴリズム[3]を用いたニューラルネットワークによる手法を提案し、その構成について述べる.

2 複素BP学習によるOFの動き解釈

我々が、二次元画像の動き成分の解釈を行った手法[4]では、複素BP学習[2]を行うニューラルネットワークを用いてOFの動きを解釈し、ネットワークの出力として、フレームに平行な移動成分(dx, dy), 拡大・縮小成分(dz), フレーム中央を中心とする回転成分(ωz)の、4つの平面上の動きを記述するパラメータが得られる.

3 三次元動き解釈のための測定系モデル

ここでは、二眼視の場合を例にとって説明する. 図1はステレオ画像法による三次元動き解釈のための測定系のモデルである. 物体座標系はX, Y, Z軸, 左右のカメラ座標系はX_l, Y_l, Z_l軸, およびX_r, Y_r, Z_r軸によりそれぞれ構成される.

ここで、2つのカメラの視野内にある移動物体の点P(x, y, z)が、両カメラ座標系において正しく対応点決定が行われるとし、求めるべき物体座標系での点Pの三次元動きパラメータは各軸方向に平行な成分(dx, dy, dz)と、各軸についての回転成分(ωx, ωy, ωz)の6つであり、一方、左右のカメラにより撮影した実画像からは、この点の各カメラフレームへの写像点について、それぞれ4つの動きパラメータ(dx_l, dy_l, dz_l, ωz_l), および(dx_r, dy_r, dz_r, ωz_r)が得られるものとする.

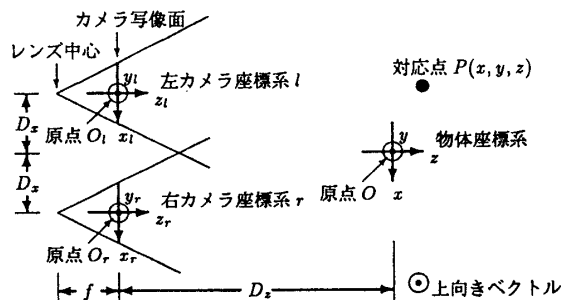


図1: ステレオ画像法による測定系モデル

4 動き成分の幾何学的変換による動き計算

図1の条件において、物体座標系における対応点Pの空間座標(x, y, z)は両カメラ座標系の視差から特定され、この座標値から左右のカメラ座標系における座標値(x_l, y_l, z_l), (x_r, y_r, z_r)に対して、およびカメラ座標系での点Pの写像点P_l, P_rと座標原点O_l, O_rとを結ぶ直線の傾斜角値(θ_l, θ_r)に対して以下の幾何変換式が成立する.

$$x_l(x, z) = \frac{D_x + x}{f + D_x + z} f \tag{1}$$

$$y_l(y, z) = \frac{y}{f + D_x + z} f \tag{2}$$

$$z_l(z) = D_x + z \tag{3}$$

$$\theta_r(x, y) = \arctan \frac{y_l}{x_l} = \arctan \frac{y}{D_x + x} \tag{4}$$

$$x_r(x, z) = \frac{x - D_x}{f + D_x + z} f \tag{5}$$

$$y_r(y, z) = \frac{y}{f + D_x + z} f \tag{6}$$

$$z_r(z) = D_x + z \tag{7}$$

$$\theta_r(x, y) = \arctan \frac{y_r}{x_r} = \arctan \frac{y}{x - D_x} \tag{8}$$

上式より、物体座標系における対応点P(x, y, z)の各動き成分(dx, dy, dz, ωx, ωy, ωz)の、左カメラ座標系への写像の各動き成分(dx_l, dy_l, dz_l, ωz_l)を計算する変換式は下式のようになる.

● dxの写像動き成分

$$\left. \begin{aligned} dx_l &= \frac{dx}{f + D_x + z} f \\ dy_l &= 0 \\ dz_l &= 0 \\ \omega z_l &= \arctan \frac{y}{D_x + x + dx} - \arctan \frac{y}{D_x + x} \end{aligned} \right) \tag{9}$$

● dyの写像動き成分

$$\left. \begin{aligned} dx_l &= 0 \\ dy_l &= \frac{dy}{f + D_x + z} f \\ dz_l &= 0 \\ \omega z_l &= \arctan \frac{y + dy}{D_x + x} - \arctan \frac{y}{D_x + x} \end{aligned} \right) \tag{10}$$

* "A proposal to interpret 3D motion of optical flow through neural network"

Akira Watanabe†, Arata Miyauchi†, Minami Miyauchi††

†武蔵工業大学, 〒158 東京都世田谷区玉堤 1-28-1

E-mail: miyauchi@cc.musashi-tech.ac.jp

††産能大学, 〒259-11 神奈川県伊勢原市上粕屋 1573

E-mail: miyauchi@miyau.mi.sanno.ac.jp

● dz の写像動き成分

$$\left. \begin{aligned} dx_i &= \frac{(D_x - x) dz}{(f + D_x + z + dz)(f + D_x + z)} f \\ dy_i &= \frac{-y dz}{(f + D_x + z + dz)(f + D_x + z)} f \\ dz_i &= dz \\ \omega_{z_i} &= 0 \end{aligned} \right) \quad (11)$$

● ω_x の写像動き成分

$$\left. \begin{aligned} dx_i &= \frac{(D_x - x)(z \cos \omega_x - y \sin \omega_x - z)}{(f + D_x + z \cos \omega_x - y \sin \omega_x)(f + D_x + z)} f \\ dy_i &= \frac{(f + D_x + z + y^2 + z^2) \sin \omega_x + (f + D_x + y)(\cos \omega_x - 1)}{(f + D_x + z \cos \omega_x - y \sin \omega_x)(f + D_x + z)} f \\ dz_i &= z \cos \omega_x - y \sin \omega_x - z \\ \omega_{z_i} &= \arctan \frac{y \cos \omega_x + z \sin \omega_x}{D_x + x} - \arctan \frac{y}{D_x + x} \end{aligned} \right) \quad (12)$$

● ω_y の写像動き成分

$$\left. \begin{aligned} dx_i &= \frac{(f + D_x + z - x^2 + z^2) \sin \omega_y + (f + D_x + x)(\cos \omega_y - 1)}{(f + D_x + z \cos \omega_y + x \sin \omega_y)(f + D_x + z)} f \\ dy_i &= \frac{-y(z \cos \omega_y + x \sin \omega_y - z)}{(f + D_x + z \cos \omega_y + x \sin \omega_y)(f + D_x + z)} f \\ dz_i &= z \cos \omega_y + x \sin \omega_y - z \\ \omega_{z_i} &= \arctan \frac{y}{D_x + x \cos \omega_y + z \sin \omega_y} - \arctan \frac{y}{D_x + x} \end{aligned} \right) \quad (13)$$

● ω_z の写像動き成分

$$\left. \begin{aligned} dx_i &= \frac{x \cos \omega_z + y \sin \omega_z - x}{f + D_x + z} f \\ dy_i &= \frac{y \cos \omega_z + x \sin \omega_z - y}{f + D_x + z} f \\ dz_i &= 0 \\ \omega_{z_i} &= \arctan \frac{y \cos \omega_z + x \sin \omega_z}{D_x + x \cos \omega_z + y \sin \omega_z} - \arctan \frac{y}{D_x + x} \end{aligned} \right) \quad (14)$$

同様にして右カメラの場合、あるいは三眼視の場合でも、物体座標系から各カメラ座標系に対する動き成分の変換式は、左カメラの場合と同様に求める事ができる。

従って、これらの変換式の逆変換式を求めて、各カメラから得た連続画像の動き解釈結果 [4] $(dx_i, dy_i, dz_i, \omega_{z_i})$ をこれに代入すれば、物体座標系における点 P の三次元動きを逆算する事が可能である。しかし、この手法では多くの計算処理を必要とし時間を要する上に、解がノイズの影響を受けやすいという欠点を持ち、実用的とはいえない。

5 三次元ベクトルニューラルネットワークによる動き解釈

複素 BP アルゴリズムを提案した新田らはまた、複素 BP アルゴリズムの自然な拡張として、多層ニューラルネットワークのすべてのユニットの結合の重みとしきい値と入出力信号を三次元ベクトルに拡張した、BP 学習アルゴリズムの三次元ベクトルバージョン (3DV-BP) [3] を提案しており、複素 BP 学習アルゴリズムが平面上の OF の解釈に有効であった [4] ように、本稿で述べているような空間的な解の解釈においても有効性が期待できる。

そこで、図 2 に示す構成の 3DV-BP アルゴリズムを応用した三次元動き解釈ニューラルネットワークを考え、このネットワークの出力として、物体座標系の 6 つの推定三次元動きパラメータを得る手法を提案する。

三次元動き解釈ニューラルネットワークは、入力層、中間層、出力層の 3 層からなり、ユニット数はそれぞれ、4 個 (または 6 個)、3 個、2 個であり、すべてのユニットの結合重みとしきい値と入出力信号は 3 つの実数からなる三次元ベクトル値を持つ。

ステレオ画像法において各カメラから得た連続画像について複素 BP 学習による OF の解釈法 [4] を用いて求めた平

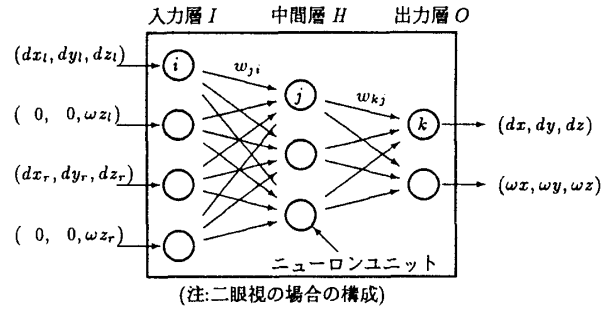


図 2: 三次元動き解釈ニューラルネットワーク

面動きパラメータ $(dx_i, dy_i, dz_i, \omega_{z_i})$ を、三次元動き解釈ニューラルネットワークの入力層のユニットに対し、それぞれ、 $(dx_i, dy_i, dz_i), (0, 0, \omega_{z_i})$ の三次元ベクトル値として入力する。

物体座標系における $(dx, dy, dz, \omega_x, \omega_y, \omega_z)$ の各基本運動パターンを個別に教師信号として与え、3DV-BP 学習を行った三次元動き解釈ニューラルネットワークに対して、上記の入力信号を与えたとき、ネットワークの出力値として 6 つの三次元動き解釈値 $(dx, dy, dz, \omega_x, \omega_y, \omega_z)$ が得られると考えられる。

6 むすび

複素 BP 学習 [2] による OF の解釈法 [4] による平面上の動きパラメータの抽出を拡張して、ステレオ画像法により二眼視で撮影した同一物体についての画像より求めた OF の解釈結果を元に、対象物体の三次元空間での動きを解釈するための手法について考察した。

物体座標系とカメラ座標系との位置関係より、幾何学的変換式を用いて、複数の OF の解釈結果の連立解により三次元動きを計算する方法は複雑な多くの計算を必要とし、実用的でない事から、OF の解釈ネットワーク [4] の出力を入力とし、三次元ベクトルバージョン BP アルゴリズム [3] を用いたニューラルネットワークによる手法を提案した。

今後の研究課題として、我々は本稿において提案した手法により、連続したステレオ実画像から、実際に物体座標系での三次元動き解釈を試みて、その能力を検証する予定である。

参考文献

- [1] 井口, 佐藤: “三次元画像計測”, 昭晃堂, (1990).
- [2] 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).
- [3] T. Nitta and Hugo de Garis: “A 3d Vector Version of the Back-propagation Algorithm”, IJCNN-92, Vol.2, pp.511-516, (1992).
- [4] M. Miyauchi, M. Seki, A. Watanabe and A. Miyauchi: “Interpretation of Optical Flow through Neural Network Learning”, IAPR, MVA'92, pp.523-528, (1992).