

3層ニューラルネットワークによる運動検出

4K-5

今村 真由美 桂井 浩

千葉工業大学情報工学科

1. はじめに

画像から対象の運動を推定することは重要であり、これまでにさまざまな研究が行われている。また、運動を検出するニューラルネットワークも多数提案されている。どのような手法を採用するにしても、最終的な運動検出に対しては、まず第1段階で独立した局所的処理により局所領域での運動を検出し、次の第2段階でそれらを統合して真の運動を求める2段階の運動検出システムのモデルが提案されている。画像に含まれる対象の運動は、その特徴の運動を把握することで検出されるという立場がある。また、局所的な運動もさまざまな大きさがあるが、これらを同じスケールで検出することが妥当であると考えられる。以上のようなことを考慮に入れて、この研究では、かなり近接した2つの時刻において得られた2枚の画像に対して、比較的小さなブロック領域におけるパターンの局所的な並進運動を検出し、最終的にはそれらを統合することで画像中の各点における運動、すなわち速度ベクトル、を求める方法について述べ、数値実験によりその動作について検討する。

2. 運動検出モジュール

運動検出モジュールは局所的な特徴の運動の速度を求める機能をもつ。このモジュールは3層のニューラルネットワークから構成される。その入力はいずれも時刻 t と $t+dt$ において得られる2枚の画像上の2つの場所におけるブロックを並置したもので与えられる。ここで、そのブロックのサイズは、運動の捕捉し易さとコストを考慮に入れて 5×5 とする。また、2つのブロックが同一の場所にあるときに、このネットワークにより検出される運動は $\langle \alpha, \beta \rangle$ 、 $-2 < \alpha, \beta < 2$ である。ただし、 α は x 方向、 β は y 方向の移動量を表す。したがって、運動の総数は25である。このような運動の例を図1に示す。



図1 運動の例

3層のニューラルネットワークから成る運動検出モジュールでは、上述のように入力ユニット数と出力ユニット数は決定されていて中間ユニット数は未定であるが、これは誤差逆伝播法を用いて上記の25通りの運動を学習させて決定する。このとき運動パターンの一般性を考慮に入れて、学習で使われる入力パターンは乱数で作成する。また、25通りの運動の学習が困難である場合には、次のようなアルゴリズムにしたがって、これらの運動を同じような数のグループに分割することとする：

- (1) 始めに、入力ユニット数50、出力ユニット数を25である1つの3層ニューラルネットワークを採用する。この場合、出力ユニットのグループは1である。
- (2) 与えられたニューラルネットワークのすべてについてそれぞれ中間ユニット数を変化させて学習を行い、いずれのネットワークにおいても学習率が所定の値を越えているかを調べる。もし、越えていたら、ここで得られているネットワークを採用して、処理を終了する。もし、所定の値を越えていない場合には、次のステップ(3)に進む。
- (3) 出力ユニットのグループを1つ増やし、その時のグループのメンバー数の差が最も少なくなるようにする。これにしたがって、ニューラルネットワークの数も1つ増加する。次にステップ(2)へ戻る。

3. 大きな運動の検出

運動検出モジュールを組み合わせて、大きな運動を検出する為のシステムを構成する。図2のように運動検出モジュールの入力領域を並べる。画像の各ブロック毎に、そのブロックにあるパターンの運動が出力されるようになっていく。 5×5 ブロックの左上隅のピクセルの位置によりブロックの位置を明示することになると、いま、プロ

Motion Detection Using 3-Layer Neural Network

Mayumi Imamura, Hiroshi Katsulai

Dept. of Computer Science, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino Chiba 275-0016, Japan

ック(i, j)に注目するとき、ブロック(i, j)と他のブロックによって構成される入力領域に対して、25通りの運動を表す出力が得られるものとする。注目しているブロック(i, j)とブロック(k, l)のに対する運動検出モジュールのn番目の出力を

$$O_{ij}^k(n)$$

で表す。このとき、ブロック(i, j)に対するこのシステムの出力は、このブロック中のパターンの運動を表しており

$$\{O_{ij}^k(n) > \tau; n=1, \dots, 25, (k, l) \in N(i, j)\}$$

となる。ただし、N(i, j)はブロック(i, j)と入力領域を構成するブロックの集合であり、 τ は許容される出力レベルを表す。いま、時刻tにおいてブロック(i, j)にあるパターンが時刻t + dtにブロック(k, l)に並進運動したとし、このときブロック(i, j)に対するモジュールが示す運動が $\langle \alpha, \beta \rangle$ であるなら、全体的な並進運動は $\langle k-i+\alpha, l-j+\beta \rangle$ で与えられる。このとき、ブロック(i, j)に対して複数の運動が検出される可能性があるため、ブロック(i, j)の近傍のブロックの出力と比較して、共通の運動出力を採用することにする。

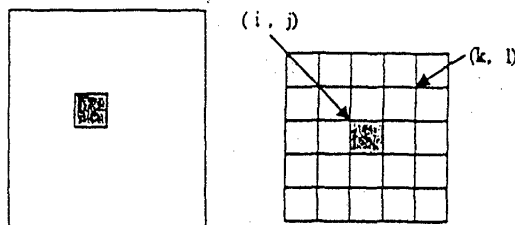


図2 運動検出モジュールの配置

4. 数値実験とその結果

まず、運動検出モジュールを決定する為に、乱数パターンを発生させて3層のニューラルネットワークの学習を行った。その結果、高い検出精度を維持する為には運動を3つにグループ分けするのが適当であることが分かった。また、この時の中間ユニット数は47であった。この決定に際して使われた、そのうち1つのネットワークの学習の過程における誤差減少の動向を図3に示す。図中のカーブは中間ユニット数に対応している。また、表1には中間ユニット数に対する誤差の標準偏差を示す。また、多数の未知のパターンに対して各運動の識別率を求めたところ未知のパターンにおけるすべての運動について表2のようになり良好な結果が得られた。

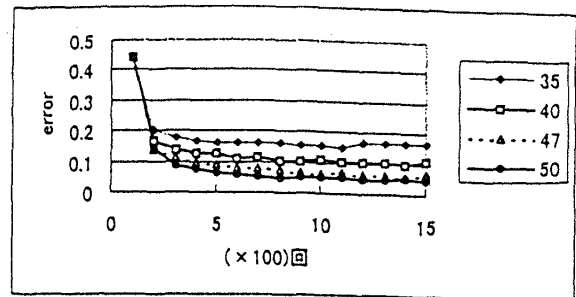


図3 ニューラルネットワークの学習経過

表1 ニューラルネットワークの学習結果

	中間ユニット数	出力				平均値	標準偏差
		$\langle -2, -2 \rangle$	$\langle -2, 0 \rangle$	$\langle 0, 0 \rangle$	$\langle 2, 0 \rangle$		
認識率 (%)	35	80.8	84.9	84.3	67.7	71.1	11.86
	40	84.4	80.4	90.7	72.7	77.2	7.85
	47	86.1	78.6	87.6	80.8	78.0	6.62
	50	89.4	87.4	86.0	83.3	81.7	6.62

表2 運動の識別率 (%)

運動	$\langle 2, 2 \rangle$	$\langle 2, 1 \rangle$	$\langle 1, 2 \rangle$	$\langle 2, 0 \rangle$	$\langle 0, 2 \rangle$	$\langle 1, 1 \rangle$	$\langle 1, 0 \rangle$
識別率	94.9	92.1	93.2	94.4	91.6	90.6	91.5
運動	$\langle 0, 1 \rangle$	$\langle 0, 0 \rangle$	$\langle 0, -1 \rangle$	$\langle -1, 0 \rangle$	$\langle 0, -2 \rangle$	$\langle -2, 0 \rangle$	
識別率	90.0	94.5	95.6	96.2	95.5	96.8	
運動	$\langle -1, -1 \rangle$	$\langle -1, -2 \rangle$	$\langle -2, -1 \rangle$	$\langle -2, -2 \rangle$	$\langle -2, 1 \rangle$	$\langle -1, 1 \rangle$	
識別率	93.9	95.4	93.9	99.2	93.9	95.5	
運動	$\langle 2, -1 \rangle$	$\langle -2, 1 \rangle$	$\langle 1, -2 \rangle$	$\langle -1, 2 \rangle$	$\langle 2, -2 \rangle$	$\langle -2, 2 \rangle$	
識別率	99.2	93.9	93.6	97.6	96.0	94.0	

さらに、大きい運動についての数値実験を行ったところ、この場合も良好な結果が得られた。

5. まとめ

本研究では、3層のニューラルネットワークから成る運動検出モジュールから構成される、運動を検出するシステムについて述べ、数値実験によりその動作を調べた。その結果、良好な検出精度を得ることができた。ただ、このシステムで大きな運動の検出を行おうとするとかなり大きな結線数が必要であり、これを低減させることが今後の課題である。

参考文献

- [1] 渥美 栄司, 横澤 一彦, 高木 幹雄: “階層的ニューラルネットによる局所運動検出”, 信学誌(D-II), J-77-D-II, pp. 606-615 (1994).
- [2] K.Nakayama: “Biological Image Motion Processing: A Review”, Vision Research, 25, 5, pp. 625-660 (1985).
- [3] S.Ullman: The Interpretation of Visual Motion, MIT Press (1979).