

具有转动不变性模式识别能力的 三层光学神经网络模型*

高成群 黄五群 沈琴婉 陈天峯

(南开大学物理系, 天津, 300071)

张延炘

(南开大学光学研究所, 天津, 300071)

摘要: 本文提出一种适于光学实现的三层神经网络模型, 该模型具有对多目标转动不变的模式识别能力. 该模型被用于对四种飞行器的平面内旋转识别. 计算机仿真表明, 该模型对样本模式、非样本模式以及有局部遮挡时, 皆有良好的联想识别能力.

关键词: 模式识别, 转动不变性, 综合鉴别函数, 光学神经网络.

引言

模式识别的不变性, 是指当目标图象千变万化时, 仍能识出是同一目标的能力. 对于复杂目标, 即使是简单的平移、旋转和尺度缩放, 用计算机进行不变性模式识别的算法也是十分复杂的, 很难在实用所要求的时间内完成.

光学信息处理的快速并行性, 使光学模式识别引起人们的兴趣. 近年来, 已提出许多具有各种不变性的光学模式识别方法^[1-5], 其中综合鉴别函数 (SDF) 方法^[1]较为引人注目. 它可将三维目标的多个投影图进行不变性编码, 从而大大压缩了信息量. 由于同一目标的任一投影皆有同样的编码, 因此对它的识别可归结为对码的识别而大大简化. SDF 的模式识别实验系统已有报道^[6], 然而该方法在存在噪声或有遮挡时, 识别能力下降, 容易产生误判.

另一方面, 人工神经网络的研究也为模式识别开拓了新的可能. 已知的许多神经网络模型^[7-10], 皆有容错和抗噪声模式识别能力; 其中 Hopfield 模型已用基于光学矢量——矩阵相乘的硬件系统的实现^[11]. 但是, 该模型不具空间不变性, 难以直接用于实际的模式识别.

本文 1990 年 9 月 28 日收到, 修改稿 1991 年 3 月 5 日收到.

* 国家自然科学基金资助项目.

本文将 SDF 方法与神经网络相结合, 提出一种既有旋转不变性, 又有容错和抗噪声能力的三层光学神经网络模型.

1 模型的构造及特点

视觉心理物理学研究表明, 人类视觉的信息处理是分层次的^[12]. 同时, 拓朴理论也证明, 用神经网络来实现任意的空间映射(异联想), 网络结构除输入层和输出层外, 至少要包含一个中间层, 即要求三层以上的网络结构^[14]. 本文提出的三层神经网络模型如图 1 所示. 外界目标成象于输入层, 产生输入层神经元的状态分布(输入模式). 输入层与中间层是一个异联想(HA)存储器, 实现对输入模式的不变性编码, 其互联可由基于前述 SDF 的原理的光学滤波技术实现. 中间层与输出层是基于 Hopfield 模型的自联想(AA)存储器, 由光学矢量矩阵相乘技术来实现. 其互联权重用光学模板在相应位置处的透过率来表示.

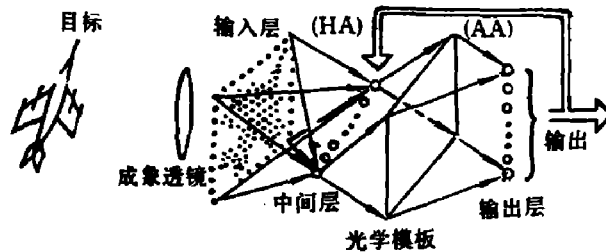


图 1 三层神经网络模型构造示意图

Fig.1 The architecture of the three-layer neural network

本模型与广泛应用的 BP 模型^[8]的不同之处在于, 它实质上是两种神经网络的级联, 即前者的输出是后者的输入. 两级网络的互联权重可分别设计(学习), 而性能却互相补足. 其优点是不必进行繁杂的学习过程, 因而也不会产生陷入局域极小而得不出稳定学习结果的困难. 更重要的是, 由于 SDF 和 Hopfield 模型的神经网络皆已成功地由光学技术实现, 因而作为两者的级联, 本模型亦将易于用光学硬件来实现.

2 原理

2.1 综合鉴别函数与异联想神经网络

前面提到, 第一级异联想网络实现目标投影图象的不变性编码, 其互联权重可由 N 个 SDF 滤波器来实现.

假定有 M 种待识别目标, 每一目标取 L 个不同的样本投影图, 这些样本用 H 个排列成方阵(本文中计算机仿真取 $H = 100 \times 100$)的神经元(这里即象元)的二值状态分布来表示(见图 2). 记为 $\xi(m, l) = [\xi_1(m, l), \xi_2(m, l), \dots, \xi_H(m, l)]$, ($m = 1, 2, \dots, M; l = 1, 2, \dots, L$) 每一个 $\xi(m, l)$ 皆可看作是 H 维二值矢量, 其分量 $\xi_h(m, l)$ 取 +1 (黑点) 或 -1 (白点).

N 个 SDF 亦由 H 维矢量 $g(n)$, ($n = 1, 2, \dots, N$) 来表示, 其分量 $g_h(n)$, ($h = 1, 2, \dots, H$)

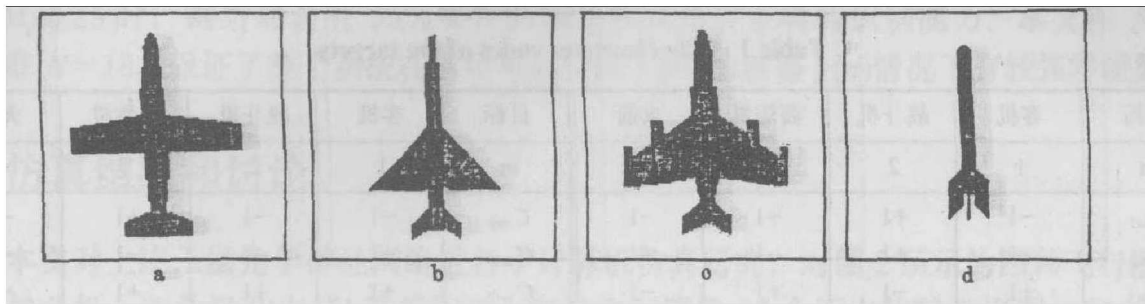


图 2 四种待识别目标未旋转的二值投影图(100×100 点阵)

(a) 客机 (b) 战斗机 (c) 轰炸机 (d) 火箭

Fig.2 The projective images of the 4 targets at zero-orientation with 100×100 pixel array

(a) Airliner (b) Fighter (c) Bomber (d) Rocket

一般不一定是二值的. 按照 SDF 原理^[1], $g(n)$ 可表示为 $\xi(m, l)$ 的线性组合:

$$g(n) = [g_1(n), g_2(n), \dots, g_H(n)] = \sum_{m=1}^M \sum_{l=1}^L a_{ml}^{(n)} \xi(m, l). \quad (1)$$

N 的大小, 由目标种类的多少和第二级神经网络的容错能力来确定.

为实现不变性编码, 要求对同一目标的所有投影样本矢量 $\xi(m, l)$, (m 一定, $l = 1, 2, \dots, L$) 与矢量 $g(n)$ 的内积为常数, 即

$$(\xi(m, l) \cdot g(n)) = C_{mn}, \quad m = 1, 2, \dots, M; n = 1, 2, \dots, N. \quad (2)$$

常数 C_{mn} 只取决于目标和 $g(n)$, 而与 l 无关, 因此 $C_{m1}, C_{m2}, \dots, C_{mN}$ 便成为目标 m 的一组不随投影样本改变的编码. 为了与下一级神经网络级联, C_{mn} 取值 +1 或 -1, 是下一级网络的输入状态.

为了求出式(1)中的系数 $a_{ml}^{(n)}$, 从而求得 $g(n)$. 我们将 $\xi(m', l')$ 与 $g(n)$ 作内积并且计及式(1)和式(2)得

$$\begin{aligned} (\xi(m', l') \cdot g(n)) &= \sum_{m=1}^M \sum_{l=1}^L a_{ml}^{(n)} (\xi(m', l') \cdot \xi(m, l)) \\ &= C_{m'n} \quad (n = 1, 2, \dots, N; m' = 1, 2, \dots, M). \end{aligned} \quad (3)$$

其中 $(\xi(m', l') \cdot \xi(m, l))$ 为第 m' 目标的第 l' 个投影矢量与第 m 目标的第 l 个投影矢量之间的内积. 当选定好编码 (如表 1 所示) 后, 求解式(3)给出的线性方程组即可得出各互联 g —矩阵的展开系数 $a_{ml}^{(n)}$, 从而得到 N 个 SDF 的 $g(n)$. 任一投影样本 $\xi(m, l)$ 与 $g(n)$ 的内积, 皆可得到对应目标的特征码 C_{mn} , 该码可用来识别目标.

由于噪声干扰的存在或非样本投影输入时, 经 HA 网络输出的码 C'_{mn} ($n = 1, 2, \dots, N$) 一般说来不严格等于 ± 1 , 而人们通常是将其取阈^[1], 即当 C'_{mn} 大于 0.5 就取为 +1, 当 C'_{mn} 小于 -0.5 就取为 -1, 这样就有可能造成误判. 为提高系统的抗干扰能力、容错能力和降低误判率, 我们将选定的特征码 C_{mn} 作为 Hopfield 模型网络的存储模式, 构成第二级自联想神经网络. 当含有错误的码 C'_{mn} 输入到该级神经网络后, 网络将通过迭代而收敛于其最接近的码 C_{mn} .

表1 目标的特征编码
Table 1 The character codes of the targets

目标	客机	战斗机	轰炸机	火箭	目标	客机	战斗机	轰炸机	火箭
m	1	2	3	4	m	1	2	3	4
C_{m1}	-1	+1	+1	-1	C_{m9}	-1	-1	+1	-1
C_{m2}	-1	+1	-1	+1	C_{m10}	+1	-1	-1	-1
C_{m3}	+1	-1	+1	-1	C_{m11}	+1	+1	+1	-1
C_{m4}	+1	-1	-1	+1	C_{m12}	+1	-1	+1	+1
C_{m5}	+1	+1	-1	-1	C_{m13}	+1	+1	-1	+1
C_{m6}	-1	-1	+1	+1	C_{m14}	-1	+1	+1	+1
C_{m7}	-1	-1	-1	+1	C_{m15}	+1	+1	+1	+1
C_{m8}	-1	+1	-1	-1	C_{m16}	-1	-1	-1	-1

2.2 目标编码 C_{mn} 的选定

在离散迭代的 Hopfield 模型中, 其互联矩阵的矩阵元为^[13~15]

$$T_{ij} = \begin{cases} \frac{1}{N} \sum_{m=1}^M C_{mi} C_{mj}, & i \neq j; \\ 0, & i = j, \end{cases} \quad (4)$$

其中 N 是神经元数目, M 是存储的模式数. $C_m = (C_{m1}, C_{m2}, \dots, C_{mN})$ 为存储模式, 且 $m = 1, 2, \dots, M$. C_{mi} 可取+1或-1分别代表神经元处于兴奋或抑制状态. 当待识别的某个存储模式为 $C_{m'}$ 时, 神经元 i 受到的局域场为

$$\begin{aligned} h_{m'i} &= \sum_{j \neq i} T_{ij} C_{m'j} \\ &= \frac{1}{N} \sum_{j \neq i} \sum_{m=1}^M C_{mi} C_{mj} C_{m'j} \approx C_{m'i} (1 + \delta_{m'i}), \end{aligned} \quad (5)$$

式中

$$\delta_{m'i} = \frac{1}{N} \sum_{j \neq i} \sum_{m \neq m'} C_{mi} C_{mj} C_{m'i} C_{m'j}, \quad (6)$$

网络迭代的动力学规则为

$$C_{mi}(t+1) = \text{Sgn}(h_{mi}(t)). \quad (7)$$

其中 t 为迭代次数, $\text{Sgn}(x)$ 表示符号函数, $\delta_{m'i}$ 是一个随机数且有如下性质:

$$\langle \delta_{m'i} \rangle = 0 \quad \text{及} \quad \langle \delta_{m'i}^2 \rangle \approx M/N.$$

由式(5)和式(7)可见, 当 $\delta_{m'i}$ 大于-1时, $C_{m'i}$ 与 $h_{m'i}$ 同号, Hopfield网络系统能实现正确识别, 否则将发生误判. 由于 $\delta_{m'i}$ 是一个随机量, 所以要改进神经网络的性能, 行之有效的办法之一是设法使 $\delta_{m'i}$ 为零, 为此我们将4组 C_{mn} 取为正交编码如表1所示. 将表1中的码值代入式(6)可得到 $\delta_{m'i} = 0$, 这显然将提高网络的模式识别能力.

码长 N 的选取(即神经元数目的选取)与网络的容错能力有关, 对一定的待识别目标数 M , N 越大, 容错能力则越强. 计算机仿真表明^[16], 在正交编码情况下, 当

$N=M/0.25$ 时, 网络对含有 50% 噪声的输入模式仍有正确的识别能力. 本文在 $M=4$ 时, 取 $N=16$, 保证了整个系统在取尽可能少的 SDF 滤波器 $g(n)$ 情况下有较强容错能力.

3 仿真结果和讨论

本文对上述三层光学神经网络进行了计算机仿真研究, 对图 2 所示的四种飞行器(客机、战斗机、轰炸机和火箭)进行识别. 每种飞行器取 36 个面内旋转投影图, 每个投影图的象元数 $H=100 \times 100$ 共 10000 个象元, 相邻两个投影图之间相隔 10° , 四种飞行器共 144 个二维投影图, 可分别记为:

$$\left. \begin{aligned} \xi(1, l_1) &= (\xi_1(1, l_1), \xi_2(1, l_1), \dots, \xi_H(1, l_1)), & (l_1 = 1, 2, \dots, 36); \\ \xi(2, l_2) &= (\xi_1(2, l_2), \xi_2(2, l_2), \dots, \xi_H(2, l_2)), & (l_2 = 1, 2, \dots, 36); \\ \xi(3, l_3) &= (\xi_1(3, l_3), \xi_2(3, l_3), \dots, \xi_H(3, l_3)), & (l_3 = 1, 2, \dots, 36); \\ \xi(4, l_4) &= (\xi_1(4, l_4), \xi_2(4, l_4), \dots, \xi_H(4, l_4)), & (l_4 = 1, 2, \dots, 36). \end{aligned} \right\} \quad (8)$$

根据上节所述算法, 首先求出对应这些目标的 16 个 SDF $g(n)$ 和由给定的 C_{mn} 码求出 T_{ij} , 以构成三层神经网络的互联. 然后对下列图形进行识别: (1) 144 个样本图形中的任意一个; (2) 面内旋转任意角度的任一非样本图形; (3) 有部分遮挡的任一样本图形. 识别方法是将上列任一图形作为输入状态, 按公式 (2) 和 (7) 进行计算机仿真. 通常第二级网络循环三次以下, 即达到稳定, 并给出识别结果, 即四种飞行器中的一种或都不是. 计算机的耗时主要是用于求 $g(n)$, 这相当于编程过程和学习过程, 当得到识别的 M 个目标不变时, 该过程是一劳永逸的. 而对目标的识别则耗时极少, 这正是实际应用所期望的.

实际三层光学神经网络将用光电混合方法来实现. 这时, 输入模式经二维光栅分成 16 个同样的图象并分别与 16 个事先制做好的 SDF 滤波器进行内积, 实现编码, 于是在中间层得到一组 16 个码的输出模式. 这组码又作为第二级网络的输入模式进行容错处理, 以鉴别其为表 1 中所列的哪一组. 很明显, 实际系统采用并行运算, 识别过程将比计算机仿真加快. 此系统目前正在研制中.

计算机仿真结果表明:

- (1) 用式(8)描述的 144 个样本图形作为输入模式, 三层网络都能正确判别目标类型.
- (2) 用四种飞行器的面内旋转任意角度的非样本模式(即上述 144 个图形以外的其它平面投影图)作为输入模式, 网络也都能正确识别. 从而实现了面内转动不变性.
- (3) 我们对各种飞行器进行部分遮挡, 当遮挡在画面的 30% 以内时, 尽管中间层输出的编码可能会出现 2~4 个误码, 但经最后一层网络容错处理后, 能使误码得以校正, 致使整个系统仍能完成目标的正确识别.
- (4) 最后我们拿上述四种飞行器之外的物体做目标输入网络, 仿真结果表明, 三层网络给出的编码与样本编码不同, 网络给出“否”的回答, 即指出目标不是要识别的四种飞行器中的任何一种.

综上所述, 本文提出了一种三层光学神经网络模型, 该模型具有面内转动不变性的模式识别能力, 同时也具有对噪声和局部遮挡的容错能力. 如果在存储模式中也包括面外转动的二维投影图, 则该模型将能推广到对三维目标的具有空间转动不变性的识别.

参考文献

- 1 Casasent D. *Appl. Opt.*, 1984, 23(10): 1620
- 2 Casasent D, Liebowitz S A. *Appl. Opt.*, 1987, 26(10): 1935
- 3 Yin Shizhuo, Cheng Lu, Mu Guoguang. *Optik*, 1989, 82(4): 129
- 4 Rieves A P et al. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1988, 10(6): 937
- 5 Schils G F, Sweeney D W. *J. Opt. Soc. Am.*, 1988, A5(8): 1309
- 6 Mu Guoguang, Cheng Lu, Yin Shizhuo. *Hybrid 3-D target recognition system based on Serial-Code Filters*, In *Proc. of Annual Meeting of OSA*, Boston, 1990,
- 7 Hopfield J J. *Neural Networks and Physical System with Emergent Collective Computational Abilities*. In *Proc. Natl. Acad. Sci. USA*. 1982, 79(7): 2554
- 8 Rumelhart D E, Hinton G E, Williams R J. *Learning Internal Representations by Error Propagation*. In *Parallel Distributed Processing*, D.E. Rumelhart and J.L. McClelland Eds., MIT Press, Cambridge, 1986, Chap. 8
- 9 Anderson J A et al. *Psychol. Rev.*, 1977, 84(5): 413
- 10 Grossberg S. *Neural Networks*, 1988, 1(1): 17
- 11 Farhat N H, Psattis D. *Appl. Opt.*, 1985, 24(10): 1469
- 12 Triesman A. *Sci. Am.*, 1986, 255(5): 106
- 13 应行仁. 模式识别与人工智能, 1990, 3(1): 29
- 14 Amit D J, Gutfreund H, Sompolinsky H. *Ann. Phys.*, 1987, 173(1): 30
- 15 Little W A, Siiaw G L. *Math. Biosci.*, 1978, 39(3/4): 281
- 16 母国光, 张延焯. 模式识别与人工智能, 1990, 3(1): 53

THREE-LAYER OPTICAL NEURAL NETWORK FOR PATTERN RECOGNITION WITH ROTATION INVARIANCE

Gao Chengqun, Huang Wuqun, Shen Qinwan, Chen Tianlun

(Department of Physics, Nankai University, Tianjin 300071, China)

Zhang Yanxin

(Institute of Modern Optics, Nankai University, Tianjin 300071, China)

Abstract: A model of three-layer optical neural network is proposed for pattern recognition with rotation invariance. As an example, the model is applied to recognize the projective views of four classes of aircrafts. Investigation of computer simulations shows that the model has good capacity of pattern recognition for patterns both inside and outside the training sets, as well as partially shaded patterns.

Key words: pattern recognition, rotation invariance, synthetic discriminant function, optical neural network.