

基于联想存储级联 WTA 模型的 旋转不变识别*

申金媛 常胜江 张延焯 母国光

南开大学光学信息技术科学开放实验室, 天津 300071
南开大学现代光学研究所, 天津 300071

摘 要 提出一种多目标旋转不变分类识别的新方法。其基本思想是多对一异联想存储模型, 为便于光学实现、充分利用系统的时间、空间带宽积, 采用了全单极(输入模式和权重均为单极)形成, 并对权重进行了二值化截取。一个分布阈值用来克服单极联想存储模型 0, 1 状态分布几率不等时存储容量急剧下降的缺点, 同时通过巧妙地选取目标所对应的异联想码将异联想模型和 WTA 模型有机地结合起来, 提高了整个系统的存储容量和容错性。本文给出了计算机仿真模拟及一个光电混合系统的实验结果。

关键词 旋转不变, 分类识别, 联想存储, WTA 模型。

1 引 言

由于多目标的旋转不变分类识别无论在军事预警、军事侦探还是在工业产品检测等方面有着广泛的应用前景, 因而引起国内外众多科学家的研究热情。他们提出了很多的方法, 如圆谐展开^[1], 串码滤波器^[2], 梅林变换^[3], 综合鉴别函数(SDFs)^[4], 锁栓滤波器(LAT)^[5], 基本辐射函数(BRF)^[6]等等。但是这些方法都存在一些缺点, 如算法复杂, 学习训练时间长, 灰度阶过大, 不便于光电硬件实现等。神经网络并行处理数据, 非常适合于解决无法用数学精确表达的问题, 而多目标的旋转不变分类识别问题正属于这类问题。

众所周知, 以往在应用联想存储模型^[7]时, 训练时做为联想的模式对都是一对一的(存储模式和输出模式联想为一对一的联想情况), 在识别时联想模式对可进行双向联想。本文将联想模式对扩展为多对一, 这样在识别时, 多个模式就可联想出一个模式。利用这个特性, 可以用联想存储模型来解决诸如多目标旋转不变分类识别, 平移不变识别, 尺度不变、手写字的认知等复杂的问题。本文以多目标的旋转不变分类识别为例, 给出计算机仿真模拟结果。为了充分利用系统的时间、空间带宽积, 采用了全单极形式(输入模式和互连权重均为单极), 并且对互连权重进行了二值化截取, 以克服光学实现多灰阶的困难。对于单极的联想存储神经网络模型, 0、1 分布几率不等会大大影响模型的存储能力和容错性, 为此提出利用一

* 国家自然科学基金、天津市二十一世纪青年基金及天津市高教局资助。

收稿日期: 1996 年 6 月 1 日; 收到修改稿日期: 1996 年 9 月 16 日

个分布阈值来弥补这一缺点。利用多对一的联想存储模型实现多目标的旋转不变性时, 一个目标的不同旋转模式都对应于一个码(称为目标的不变性编码)。如果码的位数等于目标数目, 且一个码中只有一个元素为 1, 其它为零, 那么联想存储模型输出的模式的每个元素可以认为是某一类目标的相似度, 然后通过 WTA 操作得到输出不变性编码, 这样就将联想存储模型与 WTA 模型有机地结合起来。它利用 WTA 模型的大存储容量和容错性^[8~10], 使系统的存储容量和容错性进一步提高, 进而扩大目标不变性识别的范围。

本文给出了计算机对四类飞行器进行面内旋转不变分类识别的仿真模拟, 模拟结果表明利用联想存储模型级联 WTA 模型来实现多目标的旋转不变分类是行之有效的。模拟时, 四类飞行器共有 144 个旋转模式, 取其中 72 个为学习训练模式, 另外的 72 个模式为测试模式, 正确识别率达 92.4%。文中还提出一个结构简单的光电混合实现系统(具有实时性和照明不变性), 并给出实验结果。

2 算 法

2.1 联想存储模型

联想存储模型的互连权重由相关学习规则确定^[7]:

$$W_{ijk} = \sum_{m=1}^M X_{ij}^m Y_{lk}^m, \quad (1)$$

$$i = 1, 2, \dots, I, \quad j = 1, 2, \dots, J, \quad l = 1, 2, \dots, L, \quad k = 1, 2, \dots, K$$

其中, $(\{X_{ij}^m\}, \{Y_{lk}^m\})$ 为第 m 对联想存储模式对, 它们分别为 $I \times J, L \times K$ 的二维分量, 分别称它们为 X -域和 Y -域矢量。

2.2 具有分布阈值的二值化的全单极多对一联想存储模型

在方程(1)中, 如果 $(\{X_{ij}^m\}, \{Y_{lk}^m\})$ 是双极的(元素为 1 或 -1), 识别时输入的模式也是双极的, 则模型为全双极联想存储模型; 识别时输入的模式是单极的, 则为单双极模型, 如果方程(1)中 $(\{X_{ij}^m\}, \{Y_{lk}^m\})$ 是单极的, 识别时输入的模式也是单极的, 则为全单极联想存储模型。而所谓的多对一联想存储模型, 则是指在 X -域有多个矢量与 Y -域的一个矢量对应, 或在 Y -域有多个矢量与 X -域的一个矢量对应, 显然, 多对一联想存储只能进行单向联想输出。全单极多对一联想存储模型的互连权重可由下面方程得到:

$$W_{ijk} = \sum_{n=1}^N \sum_{m=1}^M X_{ij}^{nm} Y_{lk}^m, \quad X_{ij}^{nm} = 0 \text{ 或 } 1, \quad Y_{lk}^m = 0 \text{ 或 } 1 \quad (2)$$

$$i = 1, 2, \dots, I, \quad j = 1, 2, \dots, J, \quad l = 1, 2, \dots, L, \quad k = 1, 2, \dots, K$$

其中, M 为目标的种类, N 为一类目标的模式数目。

为便于光电实现, 对(2)式所表示的互连权重进行二值化截取, 将大于 0 的权重值取为 1, 就可得到二值化的全单极多对一联想存储模型的互连权重 T_{ijk} , 联想输出为:

$$Y_{lk} = \begin{cases} 1, & (\sum_{i=1}^I \sum_{j=1}^J T_{ijk} X_{ij}^{nm}) > 0, & l = 1, 2, \dots, L, \quad k = 1, 2, \dots, K \\ 0, & (\sum_{i=1}^I \sum_{j=1}^J T_{ijk} X_{ij}^{nm}) = 0, & m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N \end{cases} \quad (3)$$

正确联想输出有 $\{Y_{lk}\} = \{Y_{lk}^m\}$ 。当存储的模式对太多时, 就会出现假态, $\{Y_{lk}\} \neq \{Y_{lk}^m\}$ 。

如果 $\{X_{ij}^{nm}\}$ 表示第 m 类目标的第 n 个旋转模式, $\{Y_{lk}^m\}$ 表示与第 m 类目标相对应的联想不变性编码, 那么利用(3)式就可以进行多目标的旋转不变分类识别。

具有分布阈值的二值化的全单极多对一联想存储模式是指进行联想识别时, 增加一个分布阈值, 即

$$Y_{lk} = \text{sgn} \left(\sum_{i=1}^I \sum_{j=1}^J T_{ijlk} X_{ij}^{nm} - \theta_{lk} \right) \quad (4)$$

$l = 1, 2, \dots, L$, $k = 1, 2, \dots, K$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$
 θ_{lk} 为分布阈值, 它与子互连矩阵中1的数目有关。分布阈值的增加克服了单极联想存储模型 0, 1 状态分布几率不等时所带来的缺点, 提高了模型的存储容量和容错性。

2.3 联想存储模型与 WTA 模型级联实现多目标的旋转不变分类识别

联想存储模型本身的存储容量和容错性并不很好, 自联想的 Hopfield 模型存储容量为 0.138, 一对一的异联想模型与自联想模型差不多, 多对一的联想模型由于只能进行单向联想不能循环迭代。所以存储容量更低, 二值化的全单极多对一联想模型由于互连权重灰度的大大压缩而进一步降低了存储容量。本文增加一个分布阈值仅仅是将单极模式中 0, 1 神经元数目分布不均所引起的存储容量及容错能力的下降得到弥补, 而不可能从根本上改变模型原有的存储容量低的缺点。因此二值化的全单极联想存储模型即使增加一个分布阈值也不可能在较大的范围内进行多目标的旋转不变分类识别。已经知道 WTA 模型有存储容量大, 容错性强等优点, 非常适合于模式识别。如果能将联想存储模型和 WTA 模型有机地结合起来, 则整个系统的存储容量和容错性可能会得到很大的提高。计算机模拟及光电实验表明本文作者的设想是正确的。

如果每个目标所对应的码 $\{Y_{lk}\}$ 中只有一个元素为 1, 其它均为 0, 则它正好是完成 WTA 操作后的输出结果, 若把(4)式右边括号中的项作为 WTA 模型中间层神经元的输入值, 经过 WTA 操作就可以得到相应的输出值, 即有:

$$Y_{lk} = \begin{cases} 1, & A_{lk} = \text{Max} (A_{lk}), & l = 1, 2, \dots, L \\ 0, & \text{其它情况}, & k = 1, 2, \dots, K \end{cases} \quad (5)$$

$$A_{lk} = \sum_{i=1}^I \sum_{j=1}^J (T_{ijlk} X_{ij}^{nm} - \theta_{lk})$$

由于 WTA 模型是一个三层网络, 所以级联 WTA 模型后还可以根据需要决定 WTA 中间层与输出层的互连权重, 再次进行异联想得到与目标有关的更多的信息, 而不仅仅是一个码。

3 计算机仿真模拟及光电混合实现系统

3.1 计算机仿真模拟

作者对四类飞行器进行了计算机仿真模拟。四类飞行器分别为轰炸机 ($\{X_{ij}^{n1}\}$)、战斗机 ($\{X_{ij}^{n2}\}$)、客机 ($\{X_{ij}^{n3}\}$)、火箭 ($\{X_{ij}^{n4}\}$), n 为旋转模式数, ij 为象素序号, $i = 1, 2, \dots, 20$, $j = 1, 2, \dots, 20$, 它们对应的不变性码分别为:

$$Y^1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Y^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y^3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad Y^4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

模拟时, 旋转模式数为 144 个(每个目标 36 个, 即每个目标每旋转 10° 取一个模式), 选 0° , 20° , 40° , \dots , 340° 的旋转模式为学习模式(共有 4×18 个), 与所对应的不变性编码通过方程(4)得到多灰阶的互连权重, 然后经过二值化截取(大于 0 的值取为 1, 等于 0 的值取为 0)得到二值的互连权重, 用所有的 144 个旋转模式分别作为输入模式进行目标的分类识别, 当模

型为全单极二值化联想存储模型时正确识别率(输出正确的模式数/所有输入模式数)为 0, 增加一个分布阈值后, 正确识别率为 40.6%, 经过级联 WTA 模型后, 正确识别率达到 92.8%。

模拟时, 涉及到分布阈值 θ_{lk} 的选取。前面已经讲到, 增加 θ_{lk} 的主要目的是克服模式中神经元状态 0, 1 分布几率不等所带来的存储容量下降的缺点, 所以它应该与模式中状态为 1 的神经元数目有关。而模式中神经元的状态直接影响到互连权重的神经元态, 因此将 θ_{lk} 分成两部分, 即 $\theta_{lk} = a_{lk} \times H_{lk}$, H_{lk} 为第 lk 个互连子矩阵中状态为 1 的神经元数目, 即:

$$H_{lk} = \sum_{i=1}^I \sum_{j=1}^J T_{ijlk}, \quad l = 1, 2, \dots, L, \quad k = 1, 2, \dots, K \quad (6)$$

显然, 得到互连权重后 H_{lk} 就已经确定, 这里 $I = 20, J = 20, L = 2, K = 2$; a_{lk} 是这样选取的: 先选一个相同的值 a , 待 a 确定后, 根据模式分类识别结果, 再在 a 的基础上选择 a_{lk} 。在计算机模拟实验中, a 的值在 0.06 ~ 0.075 之间时实验结果比较好。对于二值化的全单极联想存储模型级联 WTA 模型, 正确识别率达到 87.6%。然后根据出错的模式及互连子矩阵中神经元态为 1 的数目 H_{lk} 具体调整 a_{lk} , 进一步提高系统的正确识别率。初步调整后, 系统的正确识别率可达到 92.8%。仔细调整 a_{lk} , 系统的正确识别率有可能进一步得到提高。

3.2 光电混合实现系统

根据上述方程, 作者提出一个实现此模型的光电混合分类识别系统(图 3)。这是一个紧凑的实时模式识别系统。平面 P_1 、 P_2 和 P_3 (透镜阵列)紧密相接, 在 P_1 放置一个液晶显示阵列以显示输入模式; 在 P_2 平面放置互连模板, 将四维张量的互连权重 $\{T_{ijlk}\}$ 排列成 $L \times K$ (这里 $L = 2, K = 2$) 个互连子矩阵, 如图 2 中第一行、第二行子矩阵所示; 用平行光照明, 通过透镜阵列 (3×3) 的作用, 在透镜阵列的焦平面就可得到输入模式加权求和的光强分布; 用 CCD 器件将光强信号转换为电信号输入到计算机中, 利用计算机对信号进行分布取阈和 WTA 操作, 得到输出结果。图 1 是一个输入模式。图 2 为实验所用的互连模板。图 1、图 2 中的透明部分是为了使系统具有光照不变性而设计的(如果照明光很稳定可以略去), 这样在焦平面可以得到一个与光强有关的阈值 $\theta = I \times J \times b$, b 为平行光的光强。用计算机分布取阈时, 用 $(\theta \times a_{lk} \times H_{lk}) \div (I \times J)$ 就可以得到与光照明强度有关的分布阈值, 从而使得整个系统具有光照不变性。本文作者对所有的 144 个旋转模式都进行了分类识别, 由于硬件误差, 实验结果比计算机模拟结果差一些, 正确识别率可达到 88.9%。其中误判的 75% 是将战斗机和客机误判为轰炸机。若把战斗机去掉, 结果正确率达到 93%(计算机模拟达 98.1%)。说明这种方法与所选的目标有关。

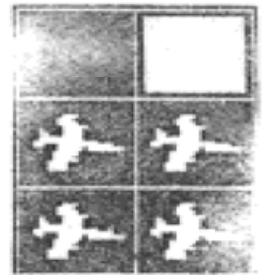


Fig. 1 One of the input pattern

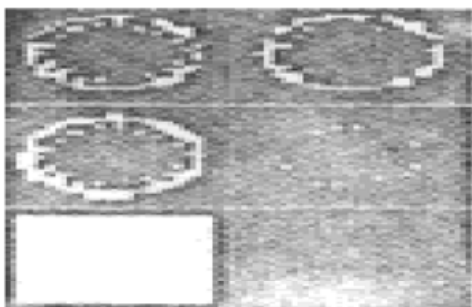


Fig. 2 Interconnection weights

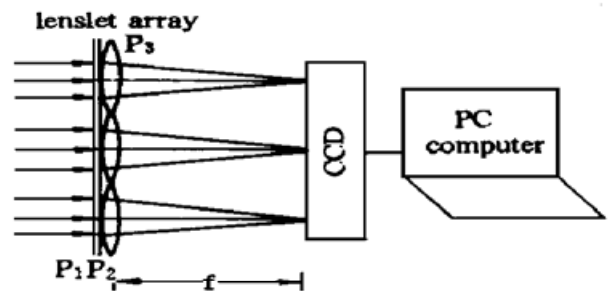


Fig. 3 The opto-electronical system

结 论 利用联想存储模型级联 WTA 模型实现多目标的旋转不变分类识别, 算法简单, 易于硬件实现, 且硬件系统紧凑, 便于集成化, 具有实时性和照明不变性等优点。

参 考 文 献

- [1] Y. N. Hsu, H. H. Aresenault, Optical pattern recognition using circular harmonic expansion. *Appl. Opt.*, 1982, **21**(4) : 4016~ 4028
- [2] S. Zh. Yin, L. Chen, G. G. Mu, 3-D target recognition by using serial-code-filters. *Optik*, 1989, **82**(4) : 129~ 131
- [3] D. Casasent, D. Psaltis, Scale invariant optical correlation using mellin transforms. *Opt. Commun.*, 1976, **17**(1) : 59~ 64
- [4] D. Casasent, Unified synthetic discrimination function computational formulation. *Appl. Opt.*, 1984, **23**(10) : 1620~ 1627
- [5] G. F. Schils, D. W. Sweeney, An optical processor for recognition of 3-D targets viewed from any direction. *J. Opt. Soc. Am. (A)*, 1988, **5**(8) : 1308~ 1321
- [6] T. Poggio, S. Edelman, A network that learns to recognize three-dimensional objects. *Nature*, 1990, **343**(1) : 263~ 266
- [7] B. Koako, Adaptive bi-directional associative memories. *Appl. Opt.*, 1987, **26**(23) : 4947~ 4960
- [8] 张延炘, 申金媛, 黄五群, 三种 WTA 模型的模拟与分析. 模式识别与人工智能, 1992, **5**(1) : 1~ 7
- [9] A. Cichocki, R. Unbehaven, Winner-take-all neural networks and its application to minimal optimization prelims. *Electron. Lett.*, 1991, **27**(22) : 2026~ 2028
- [10] J. Y. Shen, Y. X. Zhang, G. G. Mu, Implementation of modified bipolar WTA neuron network. *Chinese J. Infrared and Millimeter Waves*, 1995, **12**(2) : 245~ 252

Recognition with Rotation Invariance Based on Associative Memory Model Cascaded WTA Model

Shen Jinyuan Chang Shengjiang Zhang Yanxin Mu Guoguang
{ *Optical Information Science Laboratory, Nankai University, Tianjin 300071*
} *Institute of Modern Optics, Nankai University, Tianjin 300071*

(Received 1 June 1996; revised 16 September 1996)

Abstract A new method of multi-objects rotation invariance classification is presented. The principle of object rotation invariance classification is many-to-one associative memory model. In order to implement easily with optoelectrical techniques and to use the time-and space-bandwidth products sufficiently, unipolar model and a binary interconnection weights are used. To overcome the shortcoming of low storage capacity caused by unipolar model, a distributed thresholding is employed. Because of associate memory model cascaded WTA model by a suitable associate memory codes, the storage capacity and error-tolerant ability of the system are improved. The results of computer simulation are given and an opto-electronics system is presented.

Key words rotation invariance, classification recognition, associate memory, WTA model.