

# 一种用于多目标旋转不变分类识别的神经网络模型及算法\*

常胜江 申金媛 刘 玥 宋 庄 翟宏琛 张延焯

(南开大学现代光学研究所, 国家教委光学信息科学技术开放实验室, 天津 300071)

**摘 要** 提出了一种适于光学实现的神经网络模型和算法, 通过对3种飞行器平面内旋转投影图的识别, 证明了该方法具有对多目标转动不变的模式识别能力。

**关键词** 互连权重, 灰度阶, 适应截值。

## 1 引 言

由于目标旋转不变分类识别在工、农业产品检测及军事预警等领域有广泛的应用前景, 故对它的研究一直受到人们的广泛重视, 并已提出了多种解决方法, 如综合鉴别函数算法<sup>[1]</sup>、串码滤波器<sup>[2]</sup>、锁栓滤波器<sup>[3]</sup>以及神经网络的增量算法<sup>[4]</sup>等方法。但这些方法都存在一些缺点, 如训练时间长、灰度阶过大难以硬件实现等等。线性组合算法<sup>[5]</sup>是一种较为简单的学习算法, 它由适量的旋转样本经简单的线性叠加而成, 互连权重经二值化截值处理后网络仍具有较好的识别能力。经过分析不难看出, 由于采用模式本身构造互连权重且以内积为相似度判别, 该方法不能用于相似目标的分类识别。

与计算机模式识别相比, 光学计算与测量的精度一直是光学方法实现多目标分类识别的主要困难。二值化互连权重恰好弥补了光互连精度不高的弱点, 这种最简单的互连方式的研究一直是人们所关注的。神经网络算法中通常采用的二值化方法为 Clipping 方法, 即:

$$T_{ij} = \text{sgn}(W_{ij}) = \begin{cases} 1 & W_{ij} > 0 \\ 0 & W_{ij} = 0 \\ -1 & W_{ij} < 0 \end{cases} \quad (1)$$

其中  $W$  为截值前的互连矩阵,  $T$  为截值后的互连矩阵。这种截值方法往往会丢失互连中的一些重要信息, 使网络正确识别率下降。因此, 在保证网络基本性能的前提下, 有效地压缩互连权重的灰度阶是光学实现神经网络的关键问题之一。

本文中, 神经网络的学习算法仍采用线性组合算法, 与以往模式识别算法不同的是, 作者采用 Gram-Schmidt 正交化算法提取模式的特征信息并用这些特征信息构造网络的互连权重。为了在二值化过程中最大程度地保留正交化算法所提取的特征信息, 本文提出了一种非

\* 国家自然科学基金(No. 69587005)和南开大学自然科学基金(天津高教局)资助。

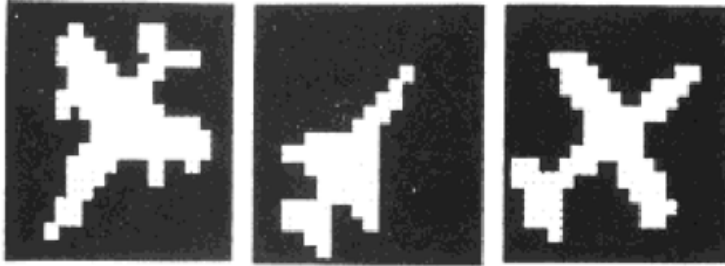
收稿日期: 1997-06-06; 收到修改稿日期: 1997-12-15

对称截值点的适应截值算法对互连权重进行了二值化处理。计算机模拟结果表明，由上述算法所建立的二值化互连神经网络模型具有较强的目标分类识别能力。

## 2 学习算法和网络结构

### 2.1 样本选取

选取了3类平面内旋转的飞行器作为待识别的目标，如图1所示。这三类飞行器为战斗



机、轰炸机和运输机，分别以  $S_F^{(i)}$ 、 $S_B^{(i)}$ 、 $S_L^{(i)}$  表示，上标  $i$  表示第  $i$  个旋转角度样本。从零方位开始，每隔  $10^\circ$  选取一个样本。每种目标可有36幅平面旋转投影图作为训练样本，3类目标共有108个样本模式。每个模式具有  $20 \times 20$  个像素，每一像素取值为0或

Fig. 1 Three kinds of aircraft as targets to be recognized 1。

### 2.2 模式的特征提取

光学中通常采用模式的边缘作为特征信息，虽然对大多数物体来说边缘包含了大部分的特征信息，但在有些情况下，物体的非边缘特征信息对模式识别来说也是非常重要的。因此，如果仅以物体的边缘作为特征信息来训练网络的互连权重，就会丢掉一些重要的非边缘特征信息，从而导致网络的性能下降。为此，在本文中采用 Gram-Schmidt 正交化算法来提取模式的特征信息，Gram-Schmidt 正交化算法是一种多维空间的投影算法。给定一组  $N$  维矢量  $S^{(m)}$ ， $m = 1, 2, \dots, M$ ， $M$  为矢量数目，正交化的矢量  $S^{*(m)}$  可写为：

$$S^{*(m)} = S^{(m)} - \sum_{l=1}^{m-1} \frac{[S^{(m)}, S^{*(l)}]}{\|S^{*(l)}\|} S^{*(l)} \tag{2}$$

$[S^{(m)}, S^{*(l)}]$  表示两者的内积， $\|S^{*(l)}\|$  是  $S^{*(l)}$  的平方模，由正交的特征可知，正交化处理后的不同矢量间的内积为零，即：

$$[S^{*(m)}, S^{*(l)}] = 0, \quad l = 1, 2, \dots, M, l \neq m \tag{3}$$

由于上述正交化算法是顺次在正交化基矢上的投影算法，因此该方法并不能保证正交化了的矢量与其他的原始矢量间的内积为零。基于上述原因，作者采用矢量间两两正交的方法，即以其中一类目标的某个方位角的投影矢量为基矢，让其他的二类目标的同方位角的投影矢量与其正交。如果以战斗机的样本集作为基矢空间，那么共有  $S_F^{(i)}$  ( $i = 1, 2, \dots, 36$ )，36个基矢量，分别让  $S_B^{(i)}$  和  $S_L^{(i)}$  与  $S_F^{(i)}$  进行正交化处理，得到  $S_{B-F}^{(i)}$  和  $S_{L-F}^{(i)}$  两组正交化的矢量。由(3)式可知：

$$[S_F^{(i)}, S_{L-F}^{(i)}] = 0, \quad [S_F^{(i)}, S_{B-F}^{(i)}] = 0 \tag{4}$$

同理，分别以  $S_L^{(i)}$  和  $S_B^{(i)}$  为基矢量时，可得正交化的二组矢量  $S_{B-L}^{(i)}$ 、 $S_{F-L}^{(i)}$  和  $S_{L-B}^{(i)}$ 、 $S_{F-B}^{(i)}$ ，上述矢量有关系式：

$$[S_B^{(i)}, S_{L-B}^{(i)}] = 0, \quad [S_B^{(i)}, S_{F-B}^{(i)}] = 0, \quad [S_L^{(i)}, S_{B-L}^{(i)}] = 0, \quad [S_L^{(i)}, S_{F-L}^{(i)}] = 0 \tag{5}$$

采用线性组合算法，分别由  $S_{B-F}^{(i)}$ 、 $S_{L-F}^{(i)}$ 、 $S_{B-L}^{(i)}$ 、 $S_{F-L}^{(i)}$ 、 $S_{L-B}^{(i)}$ 、 $S_{F-B}^{(i)}$  ( $i = 1, 2, \dots, 36$ ) 构成有六个子矩阵的互连矩阵，六个子矩阵分别表示为  $W_{B-F}$ 、 $W_{L-F}$ 、 $W_{B-L}$ 、 $W_{F-L}$ 、 $W_{L-B}$ 、 $W_{F-B}$ 。由于采用了正交化算法，在不考虑线性组合系数的符号时，由线性组合算法构造的互连矩阵是一个双极互连矩阵，它由正交化算法的特点(4)式和(5)式所决定。模式与互连子矩阵的内积的绝

对值有下列关系式:

$$\begin{aligned}
 |[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-B}]| &< |[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-B}]|, & |[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-L}]| &< |[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-L}]|, \\
 |[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-F}]| &< |[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-F}]|, & |[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-B}]| &< |[\mathbf{S}_F^{(i)}, \mathbf{W}_{F-B}]|, \\
 |[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-L}]| &< |[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-L}]|, & |[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-F}]| &< |[\mathbf{S}_L^{(i)}, \mathbf{W}_{L-F}]|, \\
 |[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-B}]| &< |[\mathbf{S}_L^{(i)}, \mathbf{W}_{L-B}]|, & |[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-L}]| &< |[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-L}]|, \\
 |[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-F}]| &< |[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-F}]|, & |[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-B}]| &< |[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-B}]|, \\
 |[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-L}]| &< |[\mathbf{S}_F^{(i)}, \mathbf{W}_{F-L}]|, & |[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-F}]| &< |[\mathbf{S}_B^{(i)}, \mathbf{W}_{B-F}]|.
 \end{aligned} \tag{6}$$

可进一步写为:

$$\begin{aligned}
 \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-B}]|} &> \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-B}]|}, & \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-L}]|} &> \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-L}]|}, \\
 \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-F}]|} &> \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-F}]|}, & \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-B}]|} &> \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{F-B}]|}, \\
 \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-L}]|} &> \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-L}]|}, & \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-F}]|} &> \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{L-F}]|}, \\
 \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-B}]|} &> \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{L-B}]|}, & \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-L}]|} &> \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{F-L}]|}, \\
 \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-F}]|} &> \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{B-F}]|}, & \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{L-B}]|} &> \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{L-B}]|}, \\
 \frac{1}{|[\mathbf{S}_L^{(i)}, \mathbf{W}_{F-L}]|} &> \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{F-L}]|}, & \frac{1}{|[\mathbf{S}_F^{(i)}, \mathbf{W}_{B-F}]|} &> \frac{1}{|[\mathbf{S}_B^{(i)}, \mathbf{W}_{B-F}]|}.
 \end{aligned} \tag{7}$$

在计算机模拟中发现, 对于所选定的三类目标, 上述(7)式是成立的。同时可以看出, 如果六个互连子矩阵按照图2的方式排列, 利用(7)式的关系表达式并采用图3的网络结构可完成对输入目标的识别。图3的结构中, 中间层由二个胜者全取(WTA)网络构成。与此网络结构对应的三类目标的不变性码分别为: 轰炸机为100100; 战斗机为010010; 运输机为001001。从图3的网络结构可以看出, 只要前三个互连子矩阵的线性组合系数为同一常数, 经非线性操作(取符号函数)后, 无论系数的正负和大小, 都不会影响输出编码的结果, 并且对后三个互连子矩阵也是一样的。这也是前文没有讨论组合系数的原因。

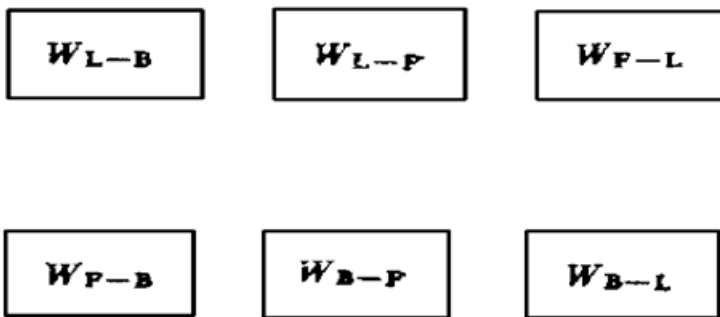


Fig. 2 Arrangement of interconnection submatrices

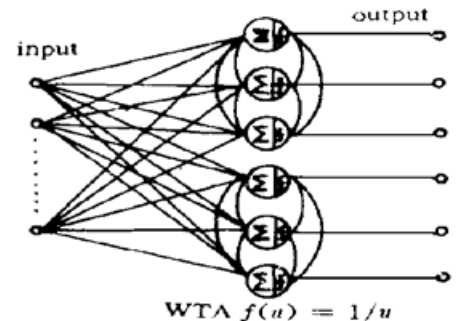


Fig. 3 The cascaded model of a neural network

### 2.3 非对称截值点的适应截值算法

由于正交化算法的引入导致上述互连矩阵  $\mathbf{W}$  的动态范围加宽, 用光学方法实现其互连权重有很大的困难, 因此互连权重的二值化是光学实现的关键问题。如何在互连矩阵的二值化过程中尽可能保存正交化算法所提取的模式特征信息(存储于互连之中) 是光学神经网络

研究的重要内容。为此作者提出了一种非对称截值点的适应截值算法对互连矩阵进行了二值化处理。由线性组合算法的特性可知,当模式为0或1状态均匀分布的随机模式时,由这些模式构成的互连矩阵  $\mathbf{W}$  统计上是一个正态分布,由于是对称的统计分布,此时截值点必定是对称的。但当模式不为随机的时,互连矩阵不再是正态分布,如果采用对称的截值算法就会破坏原互连矩阵  $\mathbf{W}$  的分布形式,导致互连权重中的信息丢失。因此,在对非随机模式条件下的互连矩阵进行二值化处理时,非对称截值点的选择是使二值化的互连权重有最佳性能的必要保证。

为了便于数学推导,把二维的图像写为一维矢量的形式,相应的互连矩阵为二维矩阵。令  $\mathbf{W}_{ij}$  为由线性组合算法得到的互连矩阵( $\mathbf{W}_{1j}$ 、 $\mathbf{W}_{2j}$ 、 $\mathbf{W}_{3j}$ 、 $\mathbf{W}_{4j}$ 、 $\mathbf{W}_{5j}$ 、 $\mathbf{W}_{6j}$ 、分别表示  $\mathbf{W}_{L-B}$ 、 $\mathbf{W}_{L-F}$ 、 $\mathbf{W}_{F-L}$ 、 $\mathbf{W}_{F-B}$ 、 $\mathbf{W}_{B-F}$ 、 $\mathbf{W}_{B-L}$ ),  $\mathbf{T}_{ij}$  为截值后的二值互连矩阵,  $\mathbf{T}_{ij}$  满足:

$$\mathbf{T}_{ij} = \begin{cases} C_1 & \mathbf{W}_{ij} > C_1 \\ 0 & -C_2 \leq \mathbf{W}_{ij} \leq C_1 \\ -C_2 & \mathbf{W}_{ij} < -C_2 \end{cases} \quad (8)$$

由截值处理引起的误差  $\Delta\mathbf{T}_{ij}$  为:

$$\Delta\mathbf{T}_{ij} = \mathbf{W}_{ij} - \mathbf{T}_{ij} = \begin{cases} \mathbf{W}_{ij} - C_1 & \mathbf{W}_{ij} > C_1 \\ \mathbf{W}_{ij} & -C_2 \leq \mathbf{W}_{ij} \leq C_1 \\ \mathbf{W}_{ij} + C_2 & \mathbf{W}_{ij} < -C_2 \end{cases} \quad (9)$$

任意一个输入模式与截值后的互连权重加权求和的模拟输出量为:

$$\begin{aligned} \mu'_i &= \sum_{j=1}^N \mathbf{T}_{ij} \mathbf{S}_j = \sum_{j=1}^N (\mathbf{W}_{ij} - \Delta\mathbf{T}_{ij}) \mathbf{S}_j = \sum_{j=1}^N \mathbf{W}_{ij} \mathbf{S}_j - \sum_{j=1}^N \Delta\mathbf{T}_{ij} \mathbf{S}_j = \\ & \mu_i - \sum_{j=1}^N \Delta\mathbf{T}_{ij} \mathbf{S}_j \end{aligned} \quad (10)$$

可进一步写为:

$$|\mu'_i - \mu_i| = \left| \sum_{j=1}^N \Delta\mathbf{T}_{ij} \mathbf{S}_j \right| \leq \sum_{j=1}^N |\Delta\mathbf{T}_{ij}| \quad (11)$$

其中  $\mu_i$  为截值前的模拟输出量,从(11)式看出,  $\mu_i$  和  $\mu'_i$  实际上是第  $i$  个互连子矩阵与输入模式的内积。为了让截值后的互连子矩阵与输入模式的内积逼近截值前的内积,最佳的截值点  $C_1^*$  和  $C_2^*$  应使  $\sum_{j=1}^N |\Delta\mathbf{T}_{ij}|$  为最小。由于互连权重为双极的,  $\sum_{j=1}^N |\Delta\mathbf{T}_{ij}|$  可写为:

$$\sum_{j=1}^N |\Delta\mathbf{T}_{ij}| = \sum_{j_1=1}^N |\Delta\mathbf{T}_{ij_1}^+| + \sum_{j_2=1}^N |\Delta\mathbf{T}_{ij_2}^-| \quad (12)$$

其中

$$\Delta\mathbf{T}_{ij_1}^+ = \begin{cases} \mathbf{W}_{ij_1} - C_1 & \mathbf{W}_{ij_1} > C_1 \\ \mathbf{W}_{ij_1} & 0 \leq \mathbf{W}_{ij_1} \leq C_1 \\ 0 & \mathbf{W}_{ij_1} < 0 \end{cases} \quad (13)$$

$$\Delta\mathbf{T}_{ij_2}^- = \begin{cases} 0 & \mathbf{W}_{ij_2} > 0 \\ \mathbf{W}_{ij_2} & -C_2 \leq \mathbf{W}_{ij_2} \leq 0 \\ \mathbf{W}_{ij_2} + C_2 & \mathbf{W}_{ij_2} < -C_2 \end{cases} \quad (14)$$

从(13)和(14)两式看出,  $\Delta\mathbf{T}_{ij_1}^+$  的值仅与  $C_1$  的取值有关而与  $C_2$  无关,而  $\Delta\mathbf{T}_{ij_2}^-$  的值仅与  $C_2$  的值有关,因此,最佳的截值点  $C_1^*$  和  $C_2^*$  应满足下面的微分方程:

$$\left. \frac{d(\sum_{j_1}^N |\Delta T_{ij_1}^+|)}{dC_1} \right|_{c_1=c_1^*} = 0 \tag{15}$$

$$\left. \frac{d(\sum_{j_2}^N |\Delta T_{ij_2}^-|)}{dC_2} \right|_{c_2=c_2^*} = 0 \tag{16}$$

从(8)式、(15)式和(16)式可见,最佳的截值点  $C_1^*$  和  $C_2^*$  是由截值前的互连矩阵  $\mathbf{W}$  决定的,而  $\mathbf{W}$  又是经样本模式训练而成,因此,  $C_1^*$  和  $C_2^*$  是由样本模式决定的二个参数。上述的截值方法是对样本模式的适应截值算法。

由于互连矩阵  $T_{ij}$  为双极的,  $T_{ij}$  可以分为  $T_{ij}^+$  和  $T_{ij}^-$  两部分:

$$T_{ij} = T_{ij}^+ - T_{ij}^- \tag{17}$$

其中  $T_{ij}^+$  是  $T_{ij}$  的正值部分,  $T_{ij}^-$  是  $T_{ij}$  的负值部分。  $T_{ij}^+$  ( $T_{ij}^-$ ) 仅由0和  $C_1^*$  ( $C_2^*$ ) 两种元素组成,因此  $T_{ij}^+$  和  $T_{ij}^-$  可写为:

$$T_{ij}^+ = C_1^* T_{ij}^{+(b)}, \quad T_{ij}^- = C_2^* T_{ij}^{-(b)}$$

其中  $T_{ij}^{+(b)}$  和  $T_{ij}^{-(b)}$  是相应的二值矩阵。因为  $C_1^*$  和  $C_2^*$  是正常数,输入模式与第  $i$  个互连子矩阵的内积  $\mu_i$  可写为:

$$\mu_i = T_{ij} S_j = (T_{ij}^+ - T_{ij}^-) S_j = C_1^* T_{ij}^{+(b)} S_j - C_2^* T_{ij}^{-(b)} S_j \tag{18}$$

在光学神经网络系统中通常采用一定的编码方法来分别表示正和负的加权求和结果。在正和负运算结果相减时,只需把正、负模板运算的结果分别乘以  $C_1^*$  和  $C_2^*$  就可以完成(18)式的运算。从硬件实现的角度来看互连矩阵实际上是一个二值化( $\pm 1, 0$ )的矩阵。

### 3 计算机模拟及光学实验结果

作者采用的神经网络硬件系统如图4所示,平行光照明互连模板(IWM),互连模板中每一互连子矩阵经透镜阵列(Lenslet array)上对应位置的小透镜及成像透镜  $L_1$  重合地成像于开关阵列(LCSA)面上,与开关阵列上的输入进行相乘,相乘的结果经透镜  $L_2$  后完成求和运算会聚于 CCD 光敏面阵列上,经 A/D 转换变为数字信号后再由计算机完成非线性操作和胜者全取运算。

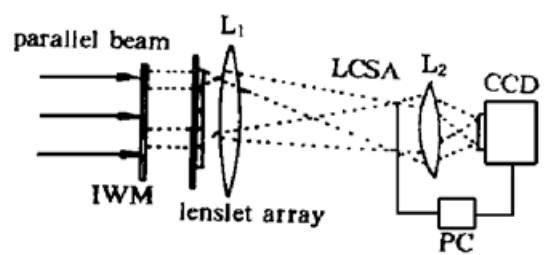


Fig. 4 Hardware system of neural network

由于上述硬件系统中透镜阵列的规模为  $32 \times 32$ , 而互连矩阵中子矩阵的个数仅为6个,因此在互连模板中把正和负互连子矩阵安排在不同的小透镜构成的子通道里,由二个小透镜完成一个互连子矩阵中正和负部分的运算,运算的结果由计算机进行相减。

采用上述的算法和实验系统,对3类飞行器进行了计算机模拟及实验研究,表1为计算机模拟及实验结果的对比。

由上面的计算机模拟和光学实验的结果可以看出,本文提出的三层神经网络模型及算法具有平面内转动不变性的模式识别能力,同时也具有对噪声的容错能力。

Table 1. Results of computer simulation and optical experiment

images to be recognized		108 training samples	108 untraining images	2×108 imagines with 30% noise
rates of recognition	computer simulation	99.1%	89.0%	92.0%
	optical experiment	97.2%	86.1%	88.7%

## 参 考 文 献

- [1] D. Casasent, Unified synthetic discrimination function computational formulation. *Appl. Opt.*, 1984, **23**(10) ·1620~ 1627
- [2] Yin Shizhuo, Cheng Lu, Mu Guoguang, 3-D target recognition by using serial-code-filter. *Optik*, 1989, **82**(4) ·129~ 131
- [3] 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) ·1309~ 1321
- [4] 高成群, 黄五群, 沈琴婉等, 具有转动不变性模式识别能力的三层光学神经网络模型. 红外与毫米波学报, 1992, **11**(1) ·69~ 75
- [5] 申金媛, 张延焯, 具有转动不变性的二值化的互连权重研究. 红外与毫米波学报, 1996, **15**(4) ·241~ 244

## A Cascaded Model of Neural Networks and Its Learning Algorithm for Pattern Recognition with Rotation Invariance

Chang Shengjiang    Shen Jinyuan    Liu Yue    Song Zhuang  
Zhai Hongchen    Zhang Yanxin

(Institute of Modern Optics, Optical Information Science Laboratory, Nankai University, Tianjin 300071)

(Received 6 June 1997; revised 15 December 1997)

**Abstract** A model of neural networks and its learning algorithm suitable for optical implementation are proposed. As an example, the model is applied to recognize the projective image of three kinds of aircraft. Computer simulations and optical experiment have shown that this model can be successfully used to error-tolerance pattern recognition of multiple targets with arbitrary orientations.

**Key words** interconnection weights, gray level, adaptive clipping.