

基于径向基函数的多目标旋转 不变分类及其光电实现*

申金媛 刘 常胜江 贾 佳 张文伟 张延 母国光

(南开大学现代光学研究所, 天津 300071)

(南开大学光学信息技术科学开放研究实验室, 天津 300071)

(1997 年 12 月 29 日收到; 1998 年 7 月 20 日收到修改稿)

提出利用径向基函数构造广义判别函数及神经网络模型, 显示了模式识别与神经网络之间的密切联系. 利用此函数和模型实现了多目标的面内(二维旋转)、面外(三维空间旋转)旋转不变识别, 计算机模拟表明, 对于旋转不变识别这是一个很有效的方法. 一个光电混合系统被提出实现此旋转不变识别方法.

PACC: 4200; 4230

1 引 言

所谓模式识别是对所研究的对象根据其共同特征和属性进行识别和分类的科学. 多目标旋转不变分类识别是模式识别中最常遇到的问题之一. 由于所解决的问题需要处理大量的信息, 因而基于冯·诺依曼原理的传统计算机存在瓶颈问题而无法快速完成这样的任务.

多目标旋转不变分类识别在军事预警、工业产品检测等方面具有广泛的应用前景, 因而引起了人们浓厚的研究兴趣. 近年来, 很多方法如圆谐展开^[1]、综合鉴别函数(SDF)^[2]、栓锁滤波器(LAT)^[3]、串码滤波器(SCF)^[4]、梅林变换^[5]、级联模型^[6]都被提出来用于多目标旋转不变分类, 但这些方法不同程度的存在诸如灰度阶过多不宜光学实现、算法复杂、存储容量小、容错力差、耗时长等缺点, 由于处理过程中信息大, 数据多, 用计算机实现不能满足现实中实时快速的要求. 人工神经网络是一种模拟人脑超分散与超平行的信息处理系统. 它不仅具有自学习和自组织功能, 而且不需要复杂的程序启动. 因而它在模式识别及图象处理中显示出相当的优越性.

神经网络与模式识别有着千丝万缕的联系, 神经网络在知识的内部表达方面有很大优势, 隐藏神经元层可抽取输入输出规则. 作为一种不明确表达的知识和经验的方法论, 它对模式识别基础理论研究有着重大的意义. 这两种方法尽管不同, 但可互相借鉴, 取长补短. 可以说, 基于模式识别聚类分析方法为神经网络处理多目标旋转不变分类识别提供了很多互连学习规则; 也可以说, 基于神经网络的模式识别分类器是模式识别领域最具有

*天津市高等教育局、天津市 21 世纪青年基金资助的课题.

吸引力的一种。

本文力图将模式识别与人工神经网络联系起来,利用基于模式识别聚类分析法构建一个网络模型.寻找一个判别函数进行聚类.如果将旋转不变模式作为一类模式,则多目标的旋转不变分类就变为模式聚类问题.本文以径向基函数(RBF)^[7]构造了一组广义判别函数,并利用它实现多目标旋转不变模式识别.文中还给出广义判别函数的神经网络解释,即利用 RBF 作为单元构建神经网络模型.由此看出,模式识别与人工神经网络间的密切联系.文献[7]证明齐次性和对称性是三层前向网络检测平移不变性和旋转不变性的充分条件,RBF 正好满足此条件,计算机模拟也表明对于多目标的旋转不变识别它是一个很好的方法.本文还给出了一个实现此模型的光电混合系统。

2 RBF 及其旋转不变识别

2.1 RBF 数学分析

在 n 维超平面中,考察一个近似函数 $f(\boldsymbol{w}, \boldsymbol{x})$,其中 \boldsymbol{w} 为常矢量, \boldsymbol{x} 为输入矢量, f 为输出矢量.给定一套训练范例 $\{X_i - f(X_i), i=1, 2, \dots, I\}$,我们的主要任务是构建函数 f 的形式,包括常矢量 \boldsymbol{w} 的确定.即找出一个分类器(classifier) f ,使该分类器能正确识别不同的模式.为易于光学实现,常矢量(又称为权值矩阵) \boldsymbol{W}_i 取 $1, 0, -1$ 三个值,本文仅取 $1, 0$ 两个值.

RBF 是一种径向对称函数^[8],本文采用高斯函数作为一种 RBF 的基元,表示为

$$f(\boldsymbol{w}, \boldsymbol{x}) = \sum_{i=1}^I \boldsymbol{w}_i \exp(-|\boldsymbol{x} - \boldsymbol{t}^i|^2 / \sigma_i^2), \quad (1)$$

其中 \boldsymbol{t}^i 为中心样本, σ_i 为间距, \boldsymbol{x} 为任意输入模式,间距

$$\sigma_i = r\rho_i = r \min(|\boldsymbol{t}^i - \boldsymbol{t}^{i+1}|, |\boldsymbol{t}^i - \boldsymbol{t}^{i-1}|), \quad (2)$$

$$\rho_i = \min(|\boldsymbol{t}^i - \boldsymbol{t}^{i+1}|, |\boldsymbol{t}^i - \boldsymbol{t}^{i-1}|, |\boldsymbol{t}^i - \boldsymbol{t}^j|) = \sqrt{\sum_{n=1}^N (\boldsymbol{t}_n^i - \boldsymbol{t}_n^j)^2} \quad (3)$$

为欧几里得距离.

如果已知 M 个目标的 I 个旋转模式,将其作为 RBF 的中心样本 $\boldsymbol{t}^i, i=1, 2, \dots, I$,任一旋转模式作为输入模式 \boldsymbol{x} ,则不同目标的 RBF 为

$$f_m(\boldsymbol{w}, \boldsymbol{x}) = \sum_{i=1}^I \boldsymbol{w}_{im} \exp(-|\boldsymbol{x} - \boldsymbol{t}^i|^2 / \sigma_i^2). \quad (4)$$

选择权值系数 \boldsymbol{w}_{im} ,使得同一类目标不同旋转模式的 RBF $f_m(\boldsymbol{w}, \boldsymbol{x})$ 具有一定的值,可以根据 $f_m(\boldsymbol{w}, \boldsymbol{x})$ 的值判别输入的模式 \boldsymbol{x} 是属于那个目标;最简单的也是最具有容错性的是选择 \boldsymbol{w}_{im} ,使得输入 \boldsymbol{W}_m 类目标的任一旋转模式, $f_m(\boldsymbol{w}, \boldsymbol{x})$ 具有最大值,实验中,我们采用了后者,因此,可以利用 RBF 进行多目标的旋转不变分类识别。

2.2 RBF 的形象解释及几何意义

用二维及三维空间图来说明方程(4)的意义.图 1 表示中心样本 \boldsymbol{t}^i ,间距 σ_i ,欲识别的

矢量 x^1, x^2 及 $x^1 - t^i$, 显然如果中心样本 t^i 选得合适, 中心样本数目 I 适当, 则在以 t^i 为中心, 以 σ_i 为半径区域内可以包括多个目标的任一旋转矢量 x . 设一个待识别的旋转模式 x 与某一中心 t^i 具有最小距离, 此最小距离小于 σ_i , 据中心样本 t^i 的类型就可以判别待识别模式的类型. 图 2 是在三维空间表示 RBF 单元 $y_i = \exp(-|x - t^i|^2 / \sigma_i^2)$, 若以 t^i 为中心原点, 则 y_i 可以作为一个包络线为高斯函数的锥体. 显然若 $|x - t^i| < \sigma_i$, 则矢量 $x - t^i$ 就落在了以 t^i 中心, σ_i 为半径的圆内, 经 e 指数运算后, y_i 落在大于 e^{-1} 的高斯包络线上. 对于多目标的旋转模式分类识别来讲, 将任意一个旋转模式作为一个输入模式, 则可以根据 y_i 在包络线上的位置确定其属于那一类目标.

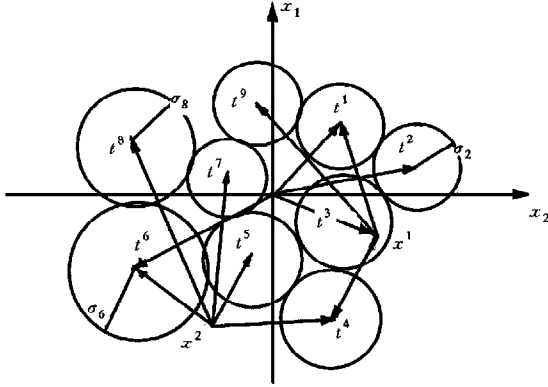


图 1 样本中心及待识别模式结构图

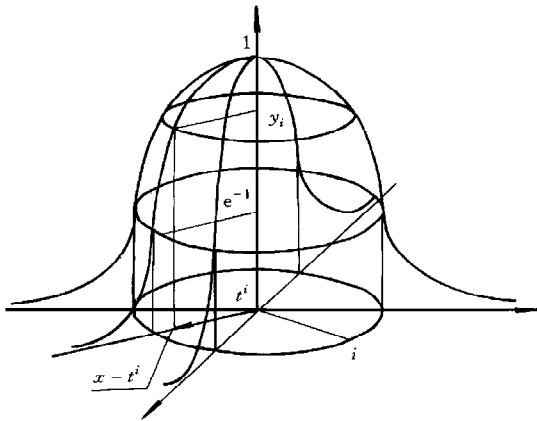


图 2 RBF 单元的几何表示

2.3 RBF 构造的广义判别函数

一个广义判别函数^[9]可以表示为

$$d(x) = \sum_{i=1}^I w_i g_i(x). \tag{5}$$

显然选择函数 $g_i(x)$ 为高斯函数时, 广义判别函数 $d(x)$ 为前面描述的 RBF 函数 $f(x)$. 利用判别函数对多类模式(设有 M 类)进行分类时, 一般考虑下面三种情况:

(a) 可用判别函数将属于 W_m 类的模式与所有其他类模式分开, 即把 M 类问题转化为 M 个两类问题. 这时需要 M 个判别函数. 其性质为

$$f_m(x) = \sum_{i=1}^l w_{im} g_i(x) \begin{cases} > 0, & x \in W_m, \\ < 0, & x \notin W_m, \end{cases} \quad m = 1, 2, \dots, M. \quad (6)$$

(b) 每一个模式类都可以与另一个模式类分开, 即每对可分, 则具有下列判别函数的形式: $f_{mj}(x) = \sum_{i=1}^l w_{mji} g_i(x)$, 这里 $f_{mj}(x) = -f_{jm}(x)$, 其判别准则为

$$\text{若 } f_{mj}(x) > 0, \quad \forall m \neq j, \quad \text{则 } x \in W_m. \quad (7)$$

(c) 对 M 类模式有 M 个判别函数

$$f_m(x) = \sum_{i=1}^l w_{mi} g_i(x), \quad m = 1, 2, \dots, M, \quad (8)$$

$$\text{若 } f_m(x) > f_j(x), \quad \forall j \neq m, \quad \text{则 } x \in W_m, \quad m, j = 1, 2, \dots, M. \quad (9)$$

本文采用情况(c), 将同一目标的不同旋转模式作为同一类模式, 选择训练样本及权值系数 w_{im} 使得输入第 W_m 类目标的任一旋转模式判别函数 $f_m(w, x)$ 都有最大值, 因而可根据 $f_m(w, x)$ 的值判断输入模式 x 的类别, 实现多目标的旋转不变分类识别.

2.4 RBF 构造的神经网络模型

如果把RBF基元定义为一个神经元, 如图3(a)所示, 则可以利用RBF构造一个神经

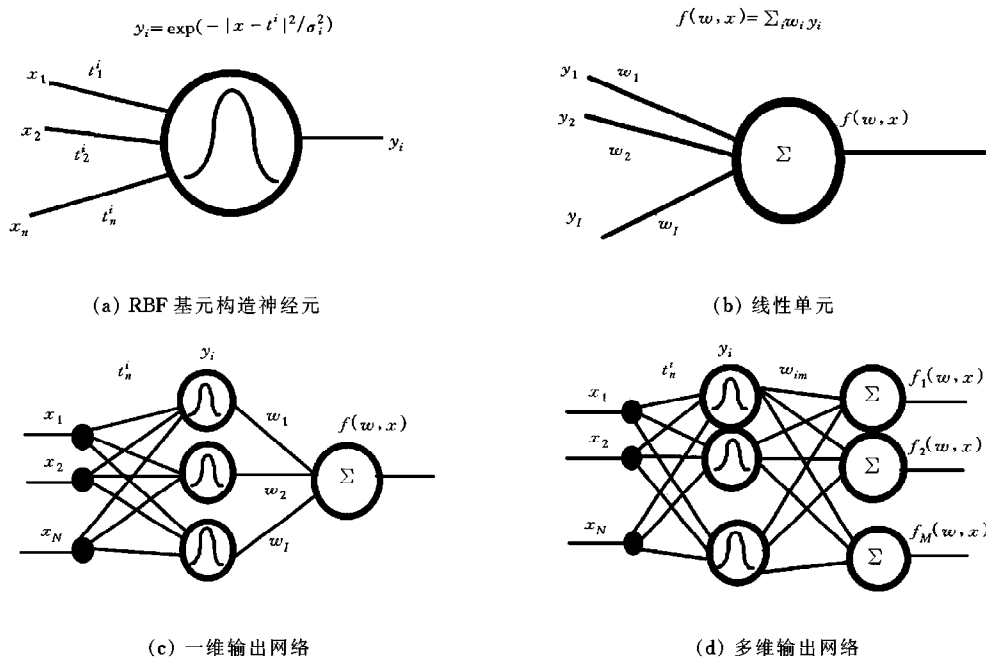


图 3 利用 RBF 构造神经网络模型

网络模型(如图 3 所示),一个 RBF 可以看作两个神经元的级联,首先是图 3(a)所示 RBF 神经元,它接受输入信息并对输入信息实现非线性操作(即高斯函数),输出结果 y_i ,然后,经过一个线性的神经元处理(如图 3(b)所示),得到输出结果 $f(\mathbf{w}, \mathbf{x})$ (如图 3(c)所示).显然,多目标的旋转不变分类识别需要一个多维输出的网络(如图 3(d)所示)来解决.如果把 RBF $f_m(\mathbf{w}, \mathbf{x})$ 作为输入模式与第 W_m 类目标间的相似度,则 RBF 构造的神经网络模型可以认为是一种 WTA 模型^[10-12].通过选择中心样本 t^i 、中心样本数目 I 和 w_{im} ,使得输入第 W_m 类目标的任一旋转模式都有最大的 $f_m(\mathbf{w}, \mathbf{x})$ 输出,据此判别输入模式 \mathbf{x} 的属类.因此利用 RBF 可以实现多目标的旋转不变分类.

3 计算机模拟实现多目标的旋转不变分类识别

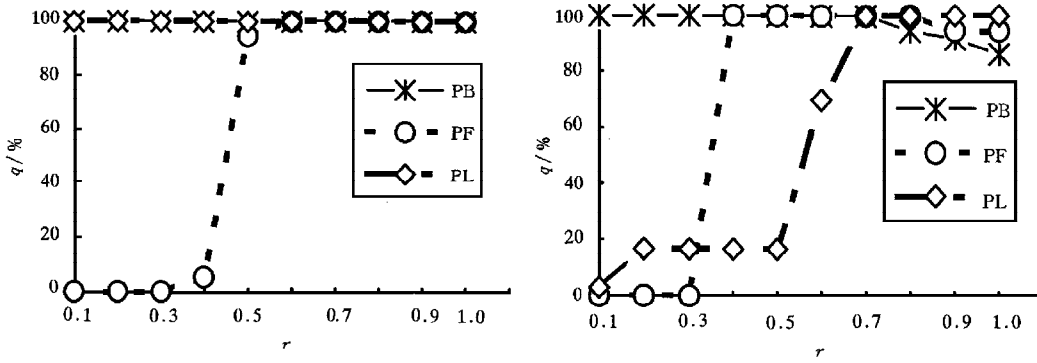
我们对三种飞机(PB:轰炸机;PF:战斗机;PL:客机)的面内、面外旋转不变识别进行了计算机模拟.

进行面内旋转不变识别模拟时,每隔 10° 取一旋转模式,三种飞机共 108 个样本.我们对三套中心样本进行了计算机模拟,第一套中心样本 t^i 为每隔 40° 的旋转模式,共有 3×9 个中心,第二套中心样本 t^i 为每隔 60° 的旋转模式,共有 3×6 个中心,第三套中心样本 t^i 为每隔 120° 的旋转模式,共有 3×3 个中心;距离为 $\sigma_i = r\rho_i$, r 为可调整的系数, ρ_i 为中心 t^i 与相邻中心的最小欧几里得距离, $\rho_i = \min(|t^i - t^{i+1}|, |t^i - t^{i-1}|)$;当 t^i 属于第 m 类目标时,权值系数 w_{im} 等于 1, 否则等于 0, 即

$$w_{im} = \begin{cases} 1 & t^i \in \Omega_m; \\ 0 & t^i \notin \Omega_m. \end{cases} \quad (10)$$

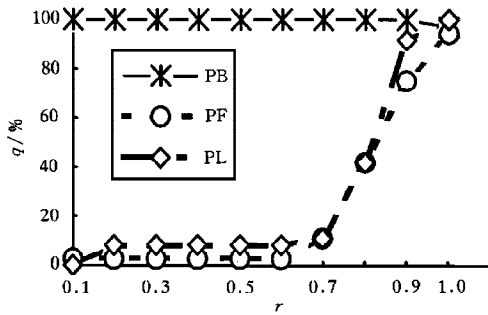
r 分别取 0.1, 0.2, ..., 1.0, 输入 108 个旋转模式进行识别,对于第一套中心样本,当 $r > 0.5$ 时,正确识别率(q)达 100%,对于第二套中心样本, $r = 0.7$ 时,正确识别率(q)达 100%,对于第三套中心样本, $r = 0.7$ 时,正确识别率达 97.2%,图 4 为输入三种飞机的旋转模式时的 r - q 曲线.

对三种飞机面外旋转(即空间旋转)进行模拟时,旋转模式及中心样本选取如下:在坐标系(x, y, z)中,令(α, β, γ)分别表示目标绕 x, y, z 轴的旋转角度,当 $\beta = 0, \gamma = 0$ 时, α 每隔 10° 每一目标选取一个投影图象作为旋转模式(面内旋转),共 $3 \times 36 = 108$ 个,选取其中 $3 \times 6 = 18$ 个模式作为中心样本(α 每隔 60°), $\beta = 10^\circ, \gamma = 0^\circ$, α 每隔 10° 每一目标选取一个投影图象作为旋转模式,共选 $3 \times 20 = 60$, ($0^\circ \leq \alpha \leq 180^\circ$),选取其中 $3 \times 4 = 12$ 个模式作为中心样本(α 每隔 60°), $\beta = 20^\circ, \gamma = 0^\circ$, α 每隔 20° 每一目标选取一个投影图象作为旋转模式,共选 $3 \times 10 = 30$, ($0^\circ \leq \alpha \leq 180^\circ$),选取其中 $3 \times 2 = 6$ 个模式作为中心样本(α 每隔 120°), $\beta = 0^\circ, \gamma = 10^\circ$, α 每隔 20° 每一目标选取一个投影图象作为旋转模式,共选 $3 \times 10 = 30$, ($0^\circ \leq \alpha \leq 180^\circ$),选取其中 $3 \times 6 = 18$ 个模式作为中心样本(α 每隔 40°);这样三种飞机共有 228 个旋转模式,其中有 $3 \times 18 = 54$ 个中心样本,权值系数 w_{im} 的选取仍如同方程(6),输入所有三维空间旋转模式进行模拟,调整系数 r 与正确识别率 q 的关系曲线如图 5 显示,当 $r > 0.7$ 时,正确识别率达到 100%.



(a) 第一套中心样本

(b) 第二套中心样本



(c) 第三套中心样本

图 4 调整系数 r 与正确识别率 q 的关系曲线

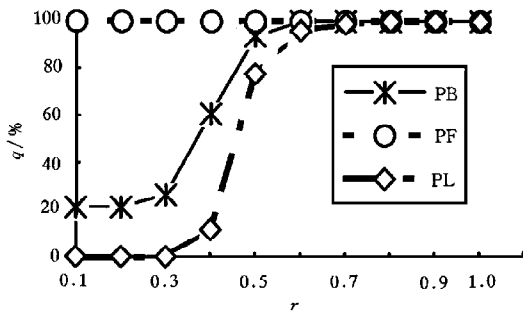


图 5 空间旋转不变识别时 r - q 曲线

4 光电实现 RBF 网络

由于中心样本和输入模式均为二值(0, 1)模式, 所以基于 RBF 的旋转不变模式识别光电混合系统易于实现. 实现时, 需要大量数据运算的间距由光学并行处理, 光学难以实

现而又神经元数目不大的 e 指数运算由计算机完成, 这样, 系统的处理速度不会受到很大的影响. 为便于光电实现, 我们将对 RBF 进行一些改写.

欧几里得距离 $|x - t^i|^2$ 改写为

$$|x - t^i|^2 = \sum_{n=1}^N (x_n \bar{t}_n^i - x_n t_n^i + t_n^i) = \sum_{n=1}^N (x_n \bar{t}_n^i - x_n t_n^i) + T^i, \text{ 式中 } T^i = \sum_{n=1}^N t_n^i, \quad (11)$$

其中 \bar{t}^i 表示中心样本 t^i 的反模式, 中心样本(训练样本)确定后, T^i 也就定了, 与输入模式无关, 据方程(4)和(11), 我们设计了一个光电混合系统实现 RBF. 光路图如图 6 所示. P_1 平面放置掩膜显示反、正中心样本模式, 紧接平面 P_1 放置一个规模大于 7×7 的透镜阵列, 在小透镜阵列的焦平面放置一个成像透镜 L_1 , 小透镜阵列和成像透镜 L_1 将所有正、反中心样本都成像在透镜 L_1 的焦平面 P_2 上, 在 P_2 平面放置实时器件(如液晶阵列)显示输入模式, 紧靠平面 P_2 放置另一成像透镜 L_2 , 经过透镜 L_2 的作用, 在平面 P_3 得到输入模式和正反中心样本的内积光强 $x't^i$, $x't^i$ (x' 表示 x 的转秩矢量), 然后通过放置在平面 P_3 的探测器阵列(PDs), 将内积光强 $x't^i$, $x't^i$ 转换为电信号, 经过 A/D 转换, 输入到计算机中, 计算机完成 $x't^i - x't^i + T^i$ (方程(11))、 e 指数运算及加权求和(方程(4))得到 $f_m(w, x)$ 的值, 并寻找最大的 $f_m(w, x)$, 判断输入模式的类型, 显示分类识别结果.

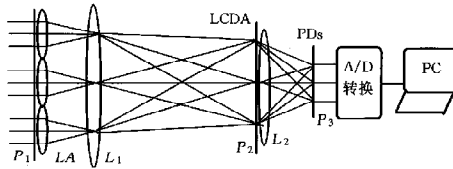


图 6 光电混合实现 RBF 识别系统

5 结论与展望

计算机模拟表明, 利用 RBF 构造广义判别函数和神经网络模型用以实现多目标的旋转不变分类识别(面内, 面外)是一个很好的方法. 作为分类识别的一个模型, 我们可以把它加以推广, 不仅仅将它用于目标旋转不变模式识别, 还可以用于视觉识别的其他方面, 只要它们在抽象元素空间具有一定相似性, 如颜色的分类, 相近颜色在频谱空间是相近的; 语音识别, 相似声音在声域空间等等, 只要我们能够找出描述这些模式的完备系作为训练样本, 就有可能利用 RBF 的方法进行识别. 可见对于模式识别来讲, 这是一个很有潜力的发展方向.

- [1] Y. N. Hsu and H. H. Arsenault, *Appl. Opt.*, **21**(1982), 4016.
- [2] D. Casasent, *Appl. Opt.*, **23**(1984), 1620.
- [3] G. F. Schils and D. W. Sweeney, *J. Opt. Soc. Am.*, **A5**(1988), 1308.
- [4] Z. Yin, L. Chen and G. G. Mu, *Optik*, **82**(4)(1989), 129.
- [5] D. Casasent and D. Psaltis, *Opt. Commun.*, **17**(1976), 59.

- [6] Y. X. Zhang *et al.*, *Opt. Comp. Proc.*, **2**(1993), 3.
- [7] Qi Xianglin *et al.*, *Proc. Nature Science*, **6**(1996), 455.
- [8] T. Poggio and S. Edelman, *Nature*, **343**(1990), 263.
- [9] K. S. Fu ed., *The Application of Pattern Recognition* (Beijing University Press, 1990).
- [10] X. Y. Zhang, J. Y. Sheng, W. Q. Huang, *Pattern Recognition & Artificial Intelligence*, **5**(1992), 1.
- [11] A. Cichocki, R. Unbehaven, *Electron. Lett.*, **27**(1991), 2026.
- [12] J. Y. Shen, Y. X. Zhang and G. G. Mu, *Chinese Journal of Infrared and Millimeter Waves*, **12**(1995), 245 (in Chinese).

CLASSIFICATION AND OPTOELECTRONICS IMPLEMENTATION OF MULTI-OBJECTS WITH ROTATION INVARIANCE BASED ON RBF FUNCTION *

SHEN JIN-YUAN LIU YUE CHANG SHENG-JIANG JIA JIA

ZHANG WEN-WEI ZHANG YAN-XIN MU GUO-GUANG

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

(*Optical Information Science Laboratory, Nankai University, Tianjin 300071*)

(Received 29 December 1997; revised manuscript received 20 July 1998)

ABSTRACT

A general discrimination function and a neural network model are constructed by using RBF function. This reveals the close relation between pattern recognition and neural network. Multi-objects classification with in-plane(2-D) and out-plane(3-D) rotation invariance is implemented by using this general discrimination function and neural network model. The computer simulation shows that it is an effective method for classifying multi-objects with rotation invariance. One of the opto-electronics recognition systems is presented to implement this function and model.

PACC: 4200; 4230

* Project supported by the Higher Education Bureau and the Foundation for 21-Century Youth of Tianjin, China.