

# 一种复合自适应分类算法

刘隆和 许俊刚 朱平云

(海军航空工程学院)

## 摘 要

本文提出一种具有监督特性的复合自适应分类算法。模拟结果表明,该算法比非自适应的直接分类算法在识别性能上有较大提高。

**关键词**——模式识别;线性分类器;聚类分析;复合分类。

## 一、引 言

在可供利用的训练样本较少的情况下,如何设计高性能的分类器是一个值得研究的课题。一般来说,训练样本越多,所得分类器的性能越好。但在许多场合,譬如制导系统中对目标的实时识别,很难得到大量的训练样本。一种很自然的想法就是在充分利用少量已获取的样本(其类别数目已知,每个样本所属类别可能未知)的同时,对每一个实时获取的样本加以利用,形成边分类、边训练的特殊解决方法。

本文就两类问题提出一种自适应复合分类算法,所列结果可推广到多类场合。

## 二、自适应复合分类系统

假设有两类目标样本随机输入,实时探测器件不断给出其样本特征矢量。系统如图1所示。聚合分类器<sup>[1]</sup>将 $N$ 个未知类别的样本自动分成两类,从 $N+1$ 开始的样本直接送给自适应分类算法。本文主要研究初级分类器训练和自适应分类算法。

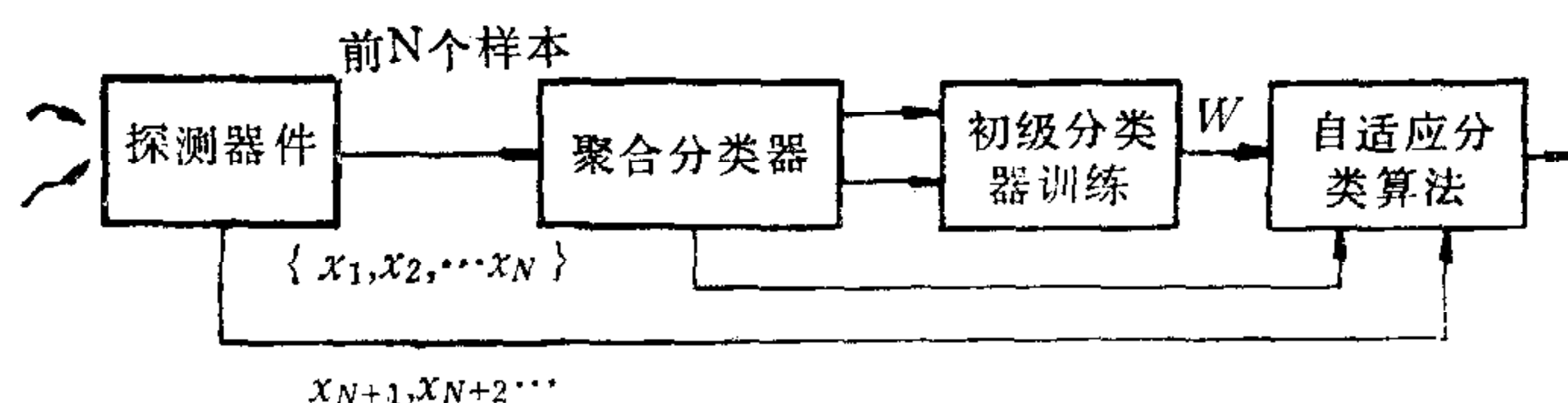


图1 自适应复合分类系统

### 三、自适应分类算法

#### 1. Robbins-Monro (RM) 算法<sup>[2]</sup>

设增广输入训练样本矢量为  $\mathbf{X} = [1, x_1, x_2, \dots, x_n]^T$ , 线性分类器权矢量为  $\mathbf{W} = [w_0, w_1, w_2, \dots, w_n]^T$ , 于是线性分类器可写为

$$\mathbf{W}^T \mathbf{X} > 0, \quad (1)$$

其中对训练样本

$$\mathbf{X} = \begin{cases} \mathbf{X}, & \text{当 } \mathbf{X} \in w_1, \\ -\mathbf{X}, & \text{当 } \mathbf{X} \in w_2. \end{cases}$$

RM 算法是一种随机逼近算法, 方法的实用性就在于即使是两类样本不可分离时也能保证算法的收敛性. 当采用一个样本修正时, 权矢量调整算法为<sup>[2]</sup>

$$\begin{aligned} \mathbf{W}_{k+1} &= \mathbf{W}_k - \rho_k \left( \frac{\partial \varepsilon^2}{\partial \mathbf{W}} \right)_k \\ &= \mathbf{W}_k - 2\rho_k (\mathbf{W}_k^T \mathbf{X}_k - d_k) \mathbf{X}_k, \end{aligned} \quad (2)$$

其中误差准则定义为

$$\varepsilon = \mathbf{W}^T \mathbf{X} - d, \quad (3)$$

$d$  为  $\mathbf{W}^T \mathbf{X}$  的期望逼近值. 收敛因子序列  $\{\rho_k\}$  必须满足 (1)

$$\lim_{k \rightarrow \infty} \rho_k = 0; \quad (2) \quad \sum_{k=1}^{\infty} \rho_k = \infty; \quad (3) \quad \sum_{k=1}^{\infty} \rho_k^2 < \infty.$$

显然,  $\rho_k$  的选择对算法的收敛性能有较大影响, 且  $\rho_k$  的选择依赖于输入样本的平均功率, 因而对 RM 算法的实际应用造成了困难. 为此, 下面提供一种  $\rho_k$  对收敛性能影响甚小的算法.

#### 2. 修正的 RM 算法 (MRM)

将 (2) 式中的修正项改为  $\left( \frac{\partial \varepsilon^2}{\partial \mathbf{W}} \right)_{k+1}$  时有下式

$$\mathbf{W}_{k+1} = \mathbf{W}_k - 2\rho_k [\mathbf{W}_{k+1}^T \mathbf{X}_{k+1} - d_{k+1}] \mathbf{X}_{k+1}. \quad (4)$$

重新整理后得

$$\mathbf{W}_{k+1} = [\mathbf{I} + 2\rho_k \mathbf{X}_{k+1} \mathbf{X}_{k+1}^T]^{-1} [\mathbf{W}_k + 2\rho_k d_{k+1} \mathbf{X}_{k+1}]. \quad (5)$$

利用矩阵求逆公式

$$[\mathbf{I} + \alpha \mathbf{Z} \mathbf{Z}^T]^{-1} = \mathbf{I} - \mathbf{Z} \mathbf{Z}^T / (\alpha^{-1} + \mathbf{Z}^T \mathbf{Z}), \quad (6)$$

上式中  $\mathbf{I}$  为单位阵. 将  $\alpha = 2\rho_k$ ,  $\mathbf{Z} = \mathbf{X}_{k+1}$  代入后合并 (5)、(6) 式并整理可得 MRM 递推算法

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\rho_k [d_{k+1} - \mathbf{W}_k^T \mathbf{X}_{k+1}] \mathbf{X}_{k+1} / (1 + 2\rho_k \mathbf{X}_{k+1}^T \mathbf{X}_{k+1}). \quad (7)$$

由上式可以看出, 当  $2\rho_k \mathbf{X}_{k+1}^T \mathbf{X}_{k+1} \gg 1$  时,  $\rho_k$  的取值对收敛速度没有影响, 实际模拟结果也证实了这一点. 本文采用 (7) 式作为训练初级分类器的算法, 并以此为基础建立了后级的自适应分类算法.

#### 3. 自适应分类算法

设聚类以后的两类样本分别为  $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_1}\}$   $\{\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_{N_2}\}$ , 样本均

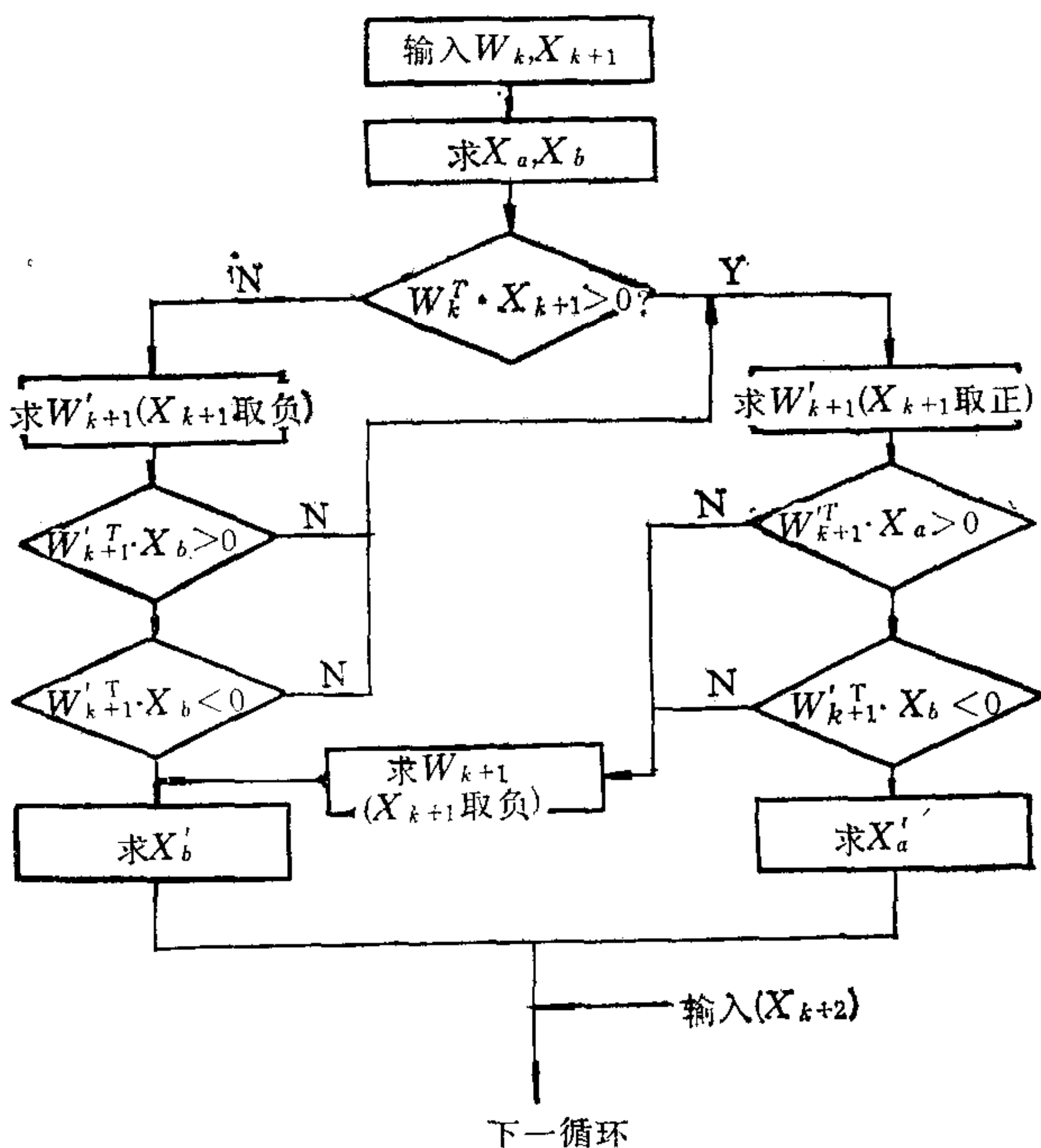


图 2 具有监督性能的自适应分类算法框图

值分别为

$$X_a = \frac{1}{N_1} \sum_{i=1}^{N_1} X_i, \quad X_i \in \omega_1, \quad (8a)$$

$$X_b = \frac{1}{N_2} \sum_{i=1}^{N_2} X'_i, \quad X'_i \in \omega_2. \quad (8b)$$

参阅图 2 所示自适应流程图。由初级分类器提供的  $W_k$  ( $W_k$  对自适应算法的影响将在下节讨论)对新输入的未知类别样本  $X_{k+1}$  进行初识别。例如设  $X_{k+1} \in \omega_1$ ，则由  $X_{k+1}$  通过 (7) 式训练出来的  $W'_{k+1}$  为可能的有效值  $W_{k+1}$ ，其可能性通过下二式监督：

$$W'_{k+1} X_a > 0, \quad (9a)$$

$$W'_{k+1} X_b < 0. \quad (9b)$$

如果 (9) 式中有一式不满足，表明  $W'_{k+1}$  比期望结果差，于是初步确定  $X_{k+1} \in \omega_2$ ，再由  $X_{k+1}$  重新训练  $W'_{k+1}$ ，这个  $W'_{k+1}$  即可直接作为有效值  $W_{k+1}$  输出，还可再进一步监督检验。不断修正  $X_a$  和  $X_b$ ，可增强监督能力，修正公式为

$$X'_a = \frac{N_1}{N_1 + 1} X_a + \frac{X_{k+1}}{N_1 + 1}, \quad \text{如 } X_{k+1} \in \omega_1, \quad (10a)$$

$$X'_b = \frac{N_2}{N_2 + 1} X_b + \frac{X_{k+1}}{N_2 + 1}, \quad \text{如 } X_{k+1} \in \omega_2. \quad (10b)$$

#### 4. 运算量考虑

由图 2 可知，完成一次分类、监督、训练的循环，可能的最大运算量是  $13n + 17$  次乘(除)法，这里  $n$  是特征矢量的维数。所以  $n$  不很大时，算法的运算量不大。

## 四、模拟结果

本文采用文献 [2] 中的标准数据  $i = 1, 2, 3$  作了计算机模拟。

### 1. 不用自适应方法而直接用 MRM 算法

先用一定数量的已知类别样本对分类器进行 MRM 训练, 然后输入检验样本观察其错误概率随训练样本数  $N$  的变化规律。图 3 中曲线 1 是数据可分离时, MRM 算法的识别性能, 取  $d_{k+1} = |W_k^T X_k|$ ,  $\rho_k = 1/(k+100)$ 。曲线 2 是数据不可分离时的情况, 由图 3 可知, 实时应用中 MRM 算法训练出来的分类器对数据不可分离场合是不能直接使用的。

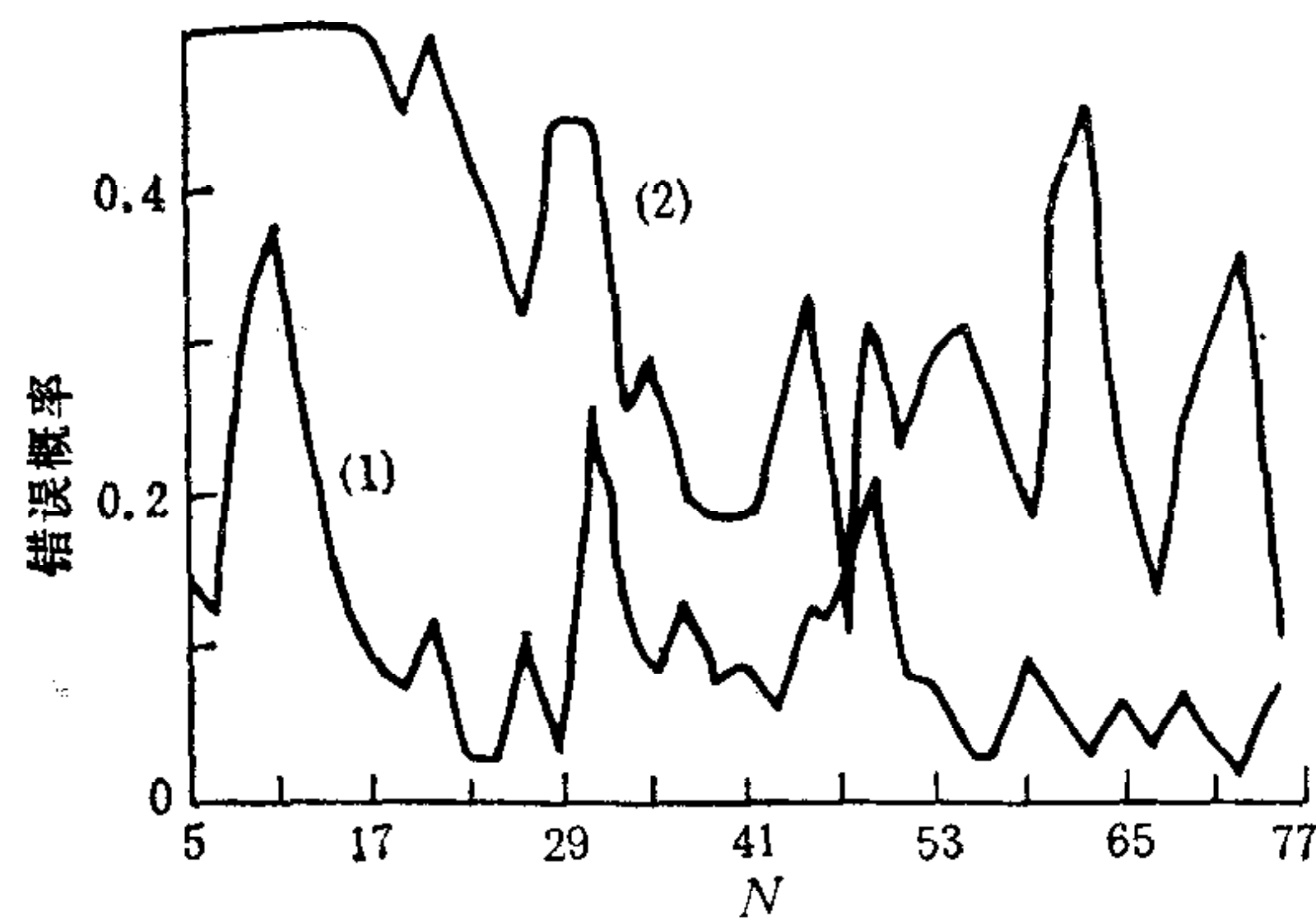


图 3 MRM 算法直接训练分类器的识别性能

### 2. 自适应分类算法的识别性能

直接给出  $\omega_1$  类 40 个样本,  $\omega_2$  类 40 个样本训练初级分类器。显然, 由初级分类器得到的  $W_k$  对自适应算法的性能产生影响, 复合自适应系统的贡献就在于即使  $W_k$  比较差 (用其引起的错误概率大小衡量), 系统仍能保持很好的分类性能。如表 1 所示, 在线性不可分离场合, 只要由初级分类器训练的  $W_k$  所引起的错误概率小于 0.38, 复合自适应分类器的错误概率就不大于 0.12。例如, 当初级线性分类器的训练样本数为 37, MRM 训练所得的  $W_k$  为

$$W_k = [0.3829, 0.1987, -0.105, -0.314, 0.551, -0.106, -0.406, -0.0616, 0.186]^T.$$

此时, 输入 100 个检验样本得其错误概率为 0.19, 而该  $W_k$  作为复合自适应分类器的初始权矢量后, 对 100 个逐个输入的未知类别样本, 错误识别率仅 0.08, 提高了识别性能。

表 1 复合自适应分类器的识别性能(线性不可分场合)

初级分类器的错误概率 $\varepsilon$	0.09	0.11	0.17	0.19	0.24	0.26	0.38
自适应分类器的错误概率 $\varepsilon$	0.09	0.07	0.12	0.08	0.12	0.12	0.40

## 参 考 文 献

[1] Tou, J. T. and Gonzales, R.C., Pattern Recognition Principles, Addison-Wesley Publishing Co-

mpany, 1974.

- [ 2 ] Fukunaga, K., Introduction to Statistical Pattern Recognition, Acad. Press. New York and London, 1972,

## A COMPOUND ADAPTIVE CLASSIFYING ALGORITHM

LIU LONGHE XU JUNGANG ZHU PINGYUN

*(Naval Aeronautical Engineering Academy)*

### ABSTRACT

A compound adaptive classifying algorithm with supervision quality is presented in this paper. The results of simulations indicate that this algorithm makes greater improvement in recognition ability than the algorithm based on non-adaptive direct classification.

**Key words** ——Pattern recognition; linear classifier; clustering analysis; compound classification.