

# 基于模糊 ART 神经网络的在线人脸识别模型的设计和实现<sup>\*</sup>)

顾 明<sup>1,2</sup>

(深圳职业技术学院软件工程系 深圳 518055)<sup>1</sup>

(哈尔滨工业大学软件工程有限公司博士后工作站 哈尔滨 150001)<sup>2</sup>

**摘 要** 本文描述了模糊 ART 神经网络的结构和特性,定义了相似函数和匹配搜索方法,通过去噪、去最小亮度和设计编码簿的方法产生人脸的特征向量图,以提取人脸特征,并用模糊 ART 神经网络对特征向量图进行识别。仿真实验证明,当选择合适的模糊 ART 神经网络参数后,该模型的在线最大识别率可以达到 81.25%,离线识别率几乎为 100%。

**关键词** 人脸识别,神经网络,模糊 ART,在线,模型

## Design and Realization of Online Face Recognition Model Based on Fuzzy ART Neural Networks

GU Ming<sup>1,2</sup>

(Dept. of Software Engineering, Shenzhen Polytechnic, Shenzhen 518055)<sup>1</sup>

(Postdoctor Workstation of Software Engineering Limited Corporation, Haerbin Industry University, Haerbin 150001)<sup>2</sup>

**Abstract** In this paper, Structure and properties of fuzzy ART are described. Choice and match function is presented. Low pass filter, cutting minimum intensity and generating vector histogram are used to extracted face features. Fuzzy ART uses the vector histogram as the original input vector data to recognize face. Simulation experiment shows that maximum online is 81.25% and offline recognition rate is nearly 100% when the fuzzy ART network parameters are selected properly.

**Keywords** Face recognition, Neural network, Fuzzy ART, Online, Model

## 1 引言

目前已经开发了不少的基于神经网络的离线人脸识别系统<sup>[1,2]</sup>,但由于这些系统只是针对事先已知的固定数量的人脸进行训练和识别,系统只能识别已知模型的人脸,并且在分类期间人脸数据库不能随时更新,因此,如果要想识别事先未知的人脸,这些系统就显得力不从心<sup>[3~5]</sup>。在线人脸识别系统能够自动地学习和存储新的人脸特征,同时,也能检测新的人脸并扩充基于新人脸的数据库。在线人脸识别所面临的主要问题是实时性和准确性,目前已有的这方面工作在解决上述这两个问题上都有一定局限,该局限来自于人脸特征的可区分性和稳定性不好,从而使得识别的准确率不高。该领域主要是以美国为主的国外在面部识别方向所做的如下一些前沿研究工作,如基于 Foirier 描述的面部轮廓识别、特征面孔的 PCA/ICA、线形区别分析、Bayesian 相似性计量、应用神经网络和基于对向量进行特征量化等。而在国内用于识别的神经网络技术主要有概率神经网络,从 AT&T 面孔数据库中和 Yale 面孔数据库中分别取样 40 人面部<sup>[6]</sup>,每人根据不同的视角、亮度和面部表情又各选 10 幅,其识别的结果为 AT&T 面孔数据库的准确率为 77.5%,Yale 面孔数据库的准确率为 73.3%。

现有的 ART1 神经网络能在二进制的环境下解决学习的稳定性及容错的同时对新的知识的敏感性问题<sup>[7,8]</sup>,但对于我们要处理的图像特征向量,要解决针对模糊量的处理问题。采用模糊集的计算方面原理,能处理模糊量的输入,但仍

需解决每次输入和系统已记忆的模式向量的近似度匹配问题。

目前采用模糊 ART 神经网络原理进行面部识别,在国内外尚属新的课题。在识别和分类的模糊 ART 神经网络中采用了在线进行学习和识别的机制,即系统对未知的输入图像可以在分类识别的同时完成修改相应权值,以应对不断到来的新的输入图像信息的处理。本文提出的模型采用仿生图像获取技术和模糊 ART 神经网络,从 AT&T 人面数据库和 Yale 人面数据库中共取 40 个人的人面图像,且每个人各取 10 个不同的面部图像进行仿真实验。创新之处在于提出一个基于人脸图像数据分析和模糊 ART(适应性共振理论)神经网络原理的在线人脸识别模型,并给出了相似函数和匹配搜索方法,其识别的结果为 AT&T 和 Yale 人面数据库的在线准确率为 78.25%,离线准确率为 99.75%,改善和提高了在线人面识别的实时性和准确性。

## 2 模糊 ART 神经网络的组成结构和特性

用模糊 ART 神经网络代替二进制的 ART 系统模型,这种替换可以把模糊集理论与 ART 系统结合起来<sup>[9~12]</sup>。在选择、搜索和学习规则中,描述 ART1 动态性的二进制操作符被模糊集理论的 AND 操作符所替代,这种替代使得 ART 系统具有更强的学习和分类能力,不但可以处理二进制的输入向量,同时也可以处理模拟输入向量。

模糊 ART 系统由三层组成,第一层  $F_0$  由  $m$  个节点组成,当前的输入向量直接通过  $F_0$  输入到整个网络中;第二层

<sup>\*</sup>)深圳市科技计划项目基金资助(05KJCD020)。顾 明 副教授,博士,主要研究方向:软件工程、神经网络。

$F_1$  也由  $m$  个节点组成,该层即接受自底向上  $F_0$  层的输入,同时也接受自顶向下  $F_2$  层的输入;第三层  $F_2$  也称为活动的分类节点。图 1 显示了模糊 ART 神经网络的组成结构。

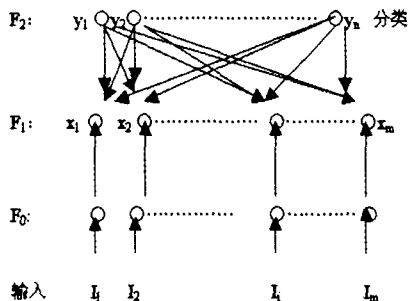


图 1 模糊 ART 神经网络的组成结构

$$F_0: I = (I_1, \dots, I_m)$$

(每个分量  $I_i$  的取值范围为  $[0 \sim 1]$  之间)

$$F_1: x = (x_1, \dots, x_m)$$

$$F_2: y = (y_1, \dots, y_n)$$

关于  $F_2$  层的每个分类节点  $j(j=1, \dots, N)$  的权重向量为:

$$W_j = (w_{j1}, \dots, w_{jm})$$

$$\text{Initially, } W_{j1}(0) = \dots = w_{jm}(0) = 1$$

$$W_j^{\text{new}} = \beta(I \wedge W_j^{\text{old}}) + (1 - \beta)W_j^{\text{old}}$$

对于快速学习,可以设置  $\beta=1$ ,上式转变为:

$$W_j^{\text{new}} = I \wedge W_j^{\text{old}}$$

$$W_j^{\text{new}} = I$$

当类  $J$  第一次成为活动节点时。

模糊 ART 的权重向量  $W_j$  即包含自底向上也包含自顶向下,模糊 ART 的动态性由选择参数  $\alpha > 0$ 、学习速率参数  $\beta \in [0, 1]$  和警戒参数  $\rho \in [0, 1]$  决定。

### 3 在线人脸识别的过程描述

#### 3.1 人脸的特征提取

首先设计一个根据图像灰度值进行特征取值的有效方案,先把一幅人面图像按  $4 \times 4$  像素分割,进行优化去噪和去掉最小亮度,再设计一个用于提取图像特征的具有 33 个特征值的编码簿(codebook),编码簿的每个特征值是  $4 \times 4$  的矩阵。将分割后的每小块图像与编码簿中的 33 个特征值进行匹配,根据出现的频度,画出一个出现概率的直方图,整个过程如图 2 所示。

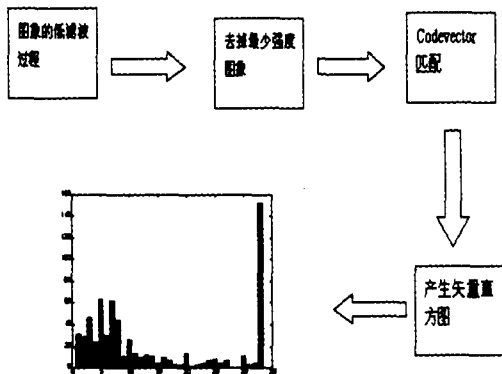


图 2 人脸特征提取的步骤

##### 3.1.1 去噪处理

去噪处理也称为低滤波,意思是当低于指定频率的信号通过时,几乎无消减,抑制高频信号的通过。通过低通过率,高频信号将被消除掉。在图像处理中,低通过率波处理可以减少高频部分并提取最有效的低频分量,提取图像的特征。

我们通过设计一个如表 1 所示的  $3 \times 3$  低通过滤波器来处理图像,过滤出来的图像然后被分成  $4 \times 4$  的具有 15/16 重叠的像素块。

表 1 去噪滤波器

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

##### 3.1.2 去最小亮度

对于图像识别而言,每个块中只有强度的变化是有意义的,为了减少每个块的 DC 偏移,寻找每个  $4 \times 4$  块中的最小强度并分别从该块中减去这个最小强度,这样一来,虽然整个图像的亮度减弱了,但亮度的变化却继续保留了下来,即图像的特征没有改变。

##### 3.1.3 产生向量柱状图

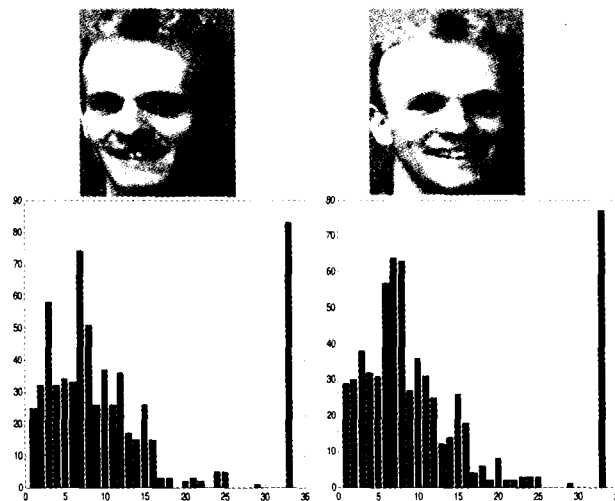


图 3 相似的柱状图

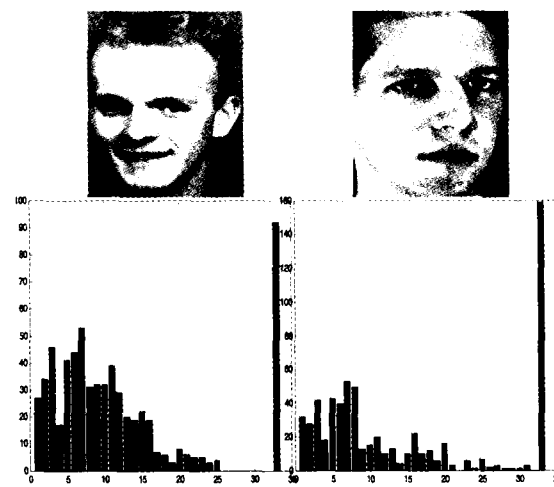


图 4 不同的柱状图

设计了一个代码簿,是具有 33 个特征值的,且每个特征值是一个  $4 \times 4$  的矩阵,从上下、左右、两个对角线,正反共 8

个方向设计强度的变化。经过了去最小亮度的 4 \* 4 的块依次作为该代码簿的输入,用 Manhattan 距离作为处理过程的匹配测量标准,通过计算输入块与代码簿之间的距离,得到输入向量的最大相似值,通过代码簿的索引号来计算输入向量和代码簿匹配的频率,得到的频率数作为产生向量柱状图的数据,这个数据反映了人脸特征的向量,产生的图称为向量柱状图,这个柱状图是由 33 个向量组成的。图 3 和图 4 显示了同一人脸变化和不同人脸变化与柱状图的对应关系。

### 3.2 模糊 ART 的识别过程

(1)对 F<sub>2</sub> 层节点的权重进行初始化,使得:

$$w_{j1(0)} = \dots = w_{jm(0)} = 1 \quad (j=1, \dots, N)$$

此时, F<sub>2</sub> 中的节点被称为是未训练过的节点,每个 w<sub>ji</sub> 是单调非递增的,因此可以收敛于某一极限值。模糊 ART 的权重向量 w<sub>ji</sub> 既包含 ART1 的自底向上,也包含 ART1 的自顶向下的权重向量。

(2)在某些 ART 系统中,当大量的输入向量作为系统的输入时,分类增值问题可能出现。通过对输入向量进行归一化处理,可以避免这个问题,归一化被认为是对输入向量的预处理。可以分  $I = a / \max(a)$  和  $I = a / \text{sum}(a)$  两种情况进行归一化处理。

(3)当输入模式 I 出现在模糊 ART 网络的输入端时,因为在初始化时 F<sub>2</sub> 层没有活动的节点,因此 F<sub>1</sub> 层的输出是  $x = I$ 。

(4)对于每个输入模式 I 和 F<sub>2</sub> 的节点 j,计算选择函数:

$$T_j(I) = \frac{|I \wedge w_j|}{a + |w_j|}$$

上式中的模糊 AND 操作符  $\wedge$  被定义为:  $((p \wedge q)_i) = \min(p_i, q_i)$ 。

操作符  $|\cdot|$  被定义为:  $|p| = \sum |p_i|, (i=1, \dots, M)$  p 和 q 是任意的 M 维向量。

通过节点  $T_j = \max\{T_j, j=1 \dots N\}$  被选中,导致分类选择被激活,此时第 j 个类被选中,使得 y<sub>j</sub> = 1,同时 y<sub>i</sub> = 0, 对其他的 j ≠ J。

(5)自顶向下在 F<sub>1</sub> 中寻找期望权重 w<sub>j</sub>。

(6)按照如下方式,使输入和 F<sub>1</sub> 中的期望权重相结合:

$x = I$ , 如果 F<sub>2</sub> 层是不活动的。

$x = I \wedge w_j$ , 如果 F<sub>2</sub> 层的第 J 个节点被选中。

(7)选择类的匹配函数定义为:

$$\frac{|I \wedge w_j|}{|I|}$$

如果警戒测试准则  $\frac{|I \wedge w_j|}{|I|} \geq \rho$  成立,则共振发生,学习也跟着发生。

如果  $\frac{|I \wedge w_j|}{|I|} < \rho$ , 则不发生共振,为了防止在搜索期间重复选择相同的类,在相同的输入模式呈现期间,设置选择函数 T<sub>j</sub> = 0, 然后返回第 3 步重新搜索一个 J 节点,直到警戒测试准则  $\frac{|I \wedge w_j|}{|I|} \geq \rho$  成立为止。

(8)一旦搜索结束,则权重向量按照下式更新权值:

$$w_j^{(new)} = \beta(I \wedge w_j^{(old)}) + (1 - \beta)w_j^{(old)}$$

这意味着权重 w<sub>j</sub><sup>(new)</sup> 从 w<sub>j</sub><sup>(old)</sup> 向 (I ∧ w<sub>j</sub><sup>(old)</sup>) 移动。

当 J 是一个未选中的节点时,设置 β = 1 是有用的,当某个类被选中后,设置 β < 1。当类 J 第一次成为活动节点时,

$$w_j^{(new)} = I。$$

(9)给出下一个输入模式,重复(1)~(8)的处理过程。

## 4 仿真实验与结果分析

### 4.1 实验环境

对由柱状图求得的 33 个向量归一化和求补后,得到 66 个向量,以此作为模糊 ART 神经网络的输入,网络没有明显的学习和分类阶段,它把这两个阶段合二为一了,因此在输出 F<sub>2</sub> 层,可以得到对输入模式的分类。

所有的仿真实验运行在 PC (Intel, Pentium(R) 4), CPU2.60GHz, 80GB 硬盘, 512MB 内存的硬件环境下,软件的操作系统用 Microsoft Windows XP Professional 2002 版本,并用 Matlab7.0 作为软件仿真环境,原始的人脸图像从 AT&T 的人面数据库和 Yale 人面数据库中共取 40 个人的面图像,且每个人各取 10 个不同的面部图像,共有 40 \* 10 = 400 个图像进行仿真实验。图像的像素点为 92 \* 112, 每人的 10 个图像包含人脸不同的角度、面部表情和光线强度的改变。

### 4.2 实验结果和分析

警戒参数值的变化从 0.7822 到 0.7906, 进行实验的警戒参数值共有 37 个,具体的警戒参数取值情况如图 5 所示。

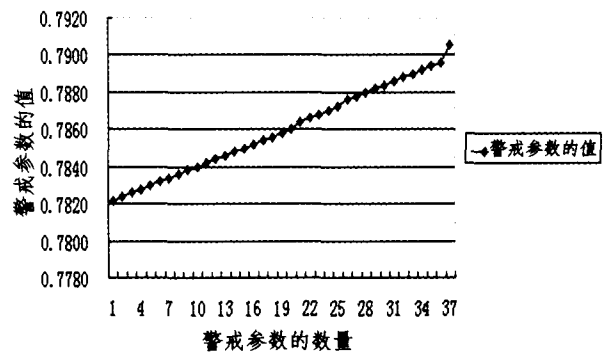


图 5 警戒值的取值情况

在线识别率定义为:正确分类的图像数量除以全部的图像数量(40 \* 10), 输入的顺序是第 1 个人的 10 个图像,第 2 个人的 10 个图像直到第 40 个人的 10 个图像,图 6 显示了警戒值和在线识别率的变化关系,该模型在警戒参数的一定取值范围内,在线识别率在 50.75% ~ 81.25% 之间,最大值为 81.25%。

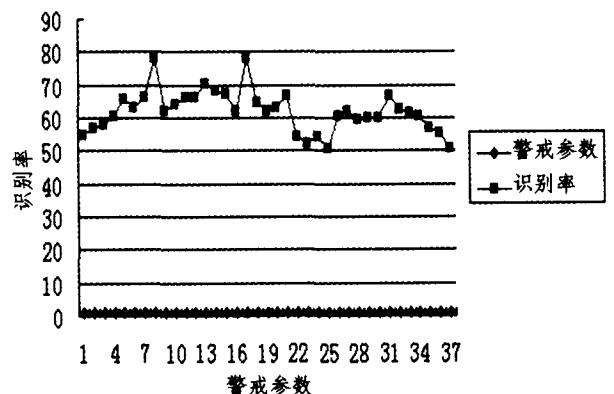


图 6 不同警戒参数值的在线识别率

图 7 数据显示的是当警戒值为 0.7856 时,前 100 个图像

的识别情况。第一行是第 1 个人的 10 个图像的识别情况，第十行是第 10 个人的 10 个图像的识别情况，如果有 100% 的在线识别率，第  $n$  行应该全为  $n$ 。

1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	4	3	3	3	3
4	2	4	4	4	5	4	4	5	5
5	5	5	5	5	5	5	5	6	6
6	3	6	6	6	6	6	6	6	6
7	6	7	6	7	7	7	2	7	7
8	8	8	1	7	8	7	8	8	8
9	9	8	9	8	9	9	9	9	5
10	9	9	10	10	9	9	10	10	10

图 7 前 10 人共 100 个图像的在线识别结果

在经过每次学习和分类阶段后，对相同的输入模式，权重向量是不同的，在警戒参数值不变的情况下，再把每个人的第 1 至第 10 个图像依次输入系统，测试识别结果，这个过程被称为离线识别率。图 8 显示了该模型在警戒参数的一定取值范围内，离线识别率在 99%~100% 之间，因此，离线识别率几乎可以达到 100%。

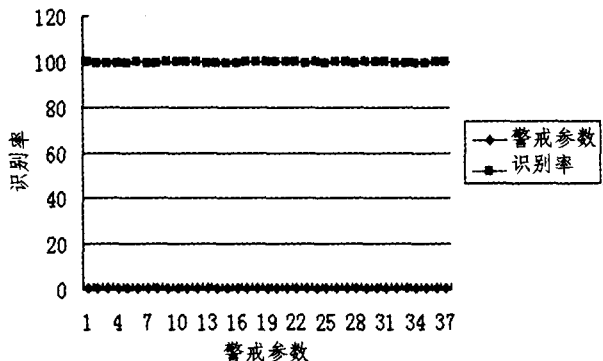


图 8 不同警戒参数值的离线识别率

图 9 数据显示的是，具有相同在线警戒值时，前 10 人的第 2、3、5、6、10 个图像的离线识别结果，比较图 9 的第一至五行与图 7 的第 2、3、5、6、10 列，发现除了 1 个不同，其余均相同，离线识别率为 99%。

1	2	3	2	5	3	6	8	9	9
1	2	3	4	5	6	7	8	5	9
1	2	3	4	5	6	7	7	8	10
1	2	4	5	5	6	7	8	9	9
1	2	3	5	6	6	7	8	5	10

图 9 前 10 人的第 2、3、5、6、10 个图像的离线识别结果

表 2 网络参数和最大识别率

学习速率 $\beta$	选择参数 $\alpha$	警戒参数	最大在线识别率 (%)	最大离线识别率 (%)	最小在线识别率 (%)	最小离线识别率 (%)	识别时间
1	0.00001	0.7836	81.25	99.5			3分 15秒 左右
1	0.00001	0.7854	81.25	99.75			
1	0.00001	0.7888				99	
1	0.00001	0.7909			50.75		

表 2 显示了模糊 ART 神经网络的网络参数同最大和最小在线和离线识别率的对应关系，对于相同的网络参数，分别有 0.7836 和 0.7854 两个警戒参数可以使网络到达最大在线

识别率，当警戒参数为 0.7888 时，网络达到最小的离线识别率，当警戒参数为 0.7909 时，网络达到最小的在线识别率。

通过实验可以得到如下结论：通过选择合适的网络参数，可以使模糊 ART 神经网络达到最大的在线和离线识别率，如表 2 的第 2 行实验数据所示。实验还表明，对每个警戒参数值，模型得到一个分类输出的平均时间为 3 分钟 15 秒左右。

原始的 ART1 系统仅仅能分类二进制的输入向量，为了处理模拟向量，我们设计并实现了通过模糊 ART 神经网络进行在线人脸识别的模型，在该模型中，用模糊子集理论的操作符替换了 ART1 中的操作符，为了解决分类增值问题，对输入向量作了归一处理。模糊 ART 神经网络属于非监督的学习，可以使学习和分类同时进行，节省了识别所需的时间。当给出认可的警戒值和一些附加的条件后，该模型可以自动判断输入的人脸是已知或未知。该模型的主要应用之一是识别已知的人脸，当检测到未知的人脸时，应用自动的在线学习机制，分类未知的人脸。

**结论和进一步的工作** 本文提出的在线人脸识别模型，在识别准确率上比起以前的方法有所提高，但还是不够理想，经过分析，发现其中一个很重要的原因是，在某些情况下，会出现一个人不同表情的两个柱状图的差别，大于某两个不同人的某种表情的差别，而人脸特征的提取是比较柱状图的差别的，这就会导致特征的提取不够准确。

进一步的研究包括，采用除了 Manhattan 距离以外的方法来测量和比较特征的相似性，同时探索和设计新的人脸特征提取方法<sup>[13,14]</sup>，研究特征提取方法和模糊 ART 的匹配适宜性，调整和改进 ART 识别模型的选择和匹配函数，调试模糊 ART 神经网络以得到最佳的网络识别参数值。

为了使该模型真正得到应用，开发相应的硬件和软件支撑平台也是今后的研究方向，包括实现模糊 ART 的硬件，固化人脸特征提取的方法，开发相应的软件支撑工具，以提高该模型的在线识别速度。

### 参考文献

- Zhao W Y, Chellappa R, Rosenfeld A, Phillips P J. Face recognition: A literature survey; [UMD CAR Technical Report CAR-TR-948]. 2000
- Turk M, Pentland A. Eigenfaces for Recognition. J. Cognitive Neuroscience, 1991, 3: 71~86
- Kotani K, Qiu C, Ohmi T. Face Recognition Using Vector Quantization Histogram Method. IEEE ICIP, Sept. 2002. 105~108
- Turk M, Pentland A. Eigenfaces for Recognition. Cognitive Neuroscience, 1991, 3: 71~86
- Patra P K, Nayak M, Nayak S K, Gobbak N K. Probabilistic neural network for pattern classification. IEEE 2002 IJCNN, May 2002. 1200~1205
- AT&T Laboratory Cambridge. The ORL Database of Faces. at <http://www.uk.research.att.com/acedatabase.html>
- Carpenter G A, Grossberg S. A massively parallel Architecture for a self-Organizing Neural Pattern Recognition Machine. Computer Vision, Graphics and Image Processing, 1987, 37:54~115
- Picton P. Neural Networks, Palgrave, 2000
- Hagan M T, Demuch H B, Beale M. Neural Network Design. PWS Publishing Company, 1996
- Carpenter G A, Grossberg S. Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps. IEEE Transactions on Neural Networks, 1992, 3(5)
- Carpenter G A. Neural Network Modes for pattern Recognition and Associative Memory. Neural Networks, 1989, 2:243~57
- Carpenter G A, Grossberg S, Rosen D B. Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System. Neural Networks, 1991, 4(4):759~771
- LI Wu-Jun, et al. A Survey of Face Recognition. Pattern Recognition and Artificial Intelligence, 2006, 19(1):58~66
- FAN Jian, WU Geng-Feng. Foremost-Policy Reinforcement Learning Based ART2 Neural Network. Pattern Recognition and Artificial Intelligence, 2006, 19(3):428~432