

基于 ART-2 神经网络及不变矩特征的图像模式识别

郝 雷, 石圣羽, 宗晓萍, 淮小利

HAO Lei, SHI Sheng-yu, ZONG Xiao-ping, HUAI Xiao-li

河北大学 电子信息工程学院, 河北 保定 071002

College of Electronic and Information Engineering, Hebei University, Baoding, Hebei 071002, China

E-mail: holly_hunter@sina.com

HAO Lei, SHI Sheng-yu, ZONG Xiao-ping, et al. Application of ART-2 neural network and invariant moment in image pattern recognition. Computer Engineering and Applications, 2007, 43(30): 44-46.

Abstract: This paper mainly uses image pre-processing and feature extraction to calculate the invariant moment of image, and ultimately realizes the image pattern recognition based on ART-2 neural network. Experimental results show that ART-2 neural network has high recognition rate. It also solves the contradiction between network's plasticity and stability, when new recognition model appears.

Key words: ART-2 neural network; feature extraction; invariant moment; pattern recognition

摘要: 将图像进行预处理并提取图像的特征, 计算出图像的不变矩, 利用 ART-2 神经网络完成了对图像的模式识别。通过实验证明 ART-2 神经网络具有较高的识别率, 并很好地解决了神经网络在模式识别中面对识别对象出现新模式时, 网络的可塑性与稳定性的矛盾。

关键词: ART-2 神经网络; 特征提取; 不变矩; 模式识别

文章编号: 1002-8331(2007)30-0044-03 **文献标识码:** A **中图分类号:** TP183

1 引言

研究模式识别的最终目的是用机器代替或部分代替人类对各种客体的辨识能力, 因此, 人工智能技术在模式识别中的应用是非常自然的。人工神经网络的自组织、自适应学习功能, 大大减弱了传统识别方法所需的约束条件, 使其对某些识别问题显示出极大的优越性^[1,2]。

目前, 用于模式识别的有代表性的神经网络主要有反向传播(BP)网络和自适应共振理论(ART)网络。BP 网络由于其结构及原理简单, 其应用最为广泛。但 BP 算法本质上属于局部寻优方法, 学习速度慢, 且容易陷入局部最小, 特别是网络的学习、记忆具有不稳定性, 使得一个训练结束的 BP 网络, 当给它提供新的记忆模式时, 将会破坏原有的记忆模式, 因此, 必须将新旧模式同时提供给网络重新进行训练, 直到达到新的稳定状态^[3]。本文采用 ART-2 神经网络进行模式识别, 很好地解决了面对新模式, 网络的稳定性问题。

2 图像特征提取及不变矩

2.1 图像模式识别中的特征提取

特征提取是对模式所包含的输入信息进行处理和分析, 将不易受随机因素干扰的信息作为该模式的特征提取出来。特征提取过程是去除冗余信息的过程, 具有提高识别精度、减少运算量和提高运算速度的作用。良好的特征应具有可区分性、稳

定性、独立性以及数目小等特点。特征提取是图像模式识别的基础。

2.2 图像的不变矩

矩是一种非常重要的表示目标总体形状的特征量。二维图形的几个关键特征均直接与矩有关, 如图像的大小、形心等。不变矩由于概念清晰, 识别率稳定, 对具有旋转和缩放变化的目标有良好的不变性及抗干扰性, 能有效地反映图像的本质特征^[4]。

设一幅 $m \times n$ 的二维离散图像灰度函数用 $f(x, y)$ 表示, 其 $p+q$ 阶矩定义为

$$m_{pq} = \sum_{x=1}^m \sum_{y=1}^n x^p y^q f(x, y)$$

两个一阶矩(m_{10}, m_{01})表征了图像的灰度重心位置, 当摄像机与目标之间发生平行于图像平面的相对运动时, 重心位置随之发生变化。

图像的 $p+q$ 阶中心矩定义为

$$\mu_{pq} = \sum_{x=1}^m \sum_{y=1}^n (x-x_0)^p (y-y_0)^q f(x, y)$$

其中 $x_0 = m_{10}/m_{00}, y_0 = m_{01}/m_{00}$ 表示图像的灰度重心坐标。由于中心矩仅具有平移不变性, 为了得到具有伸缩不变性的矩, 定义归一化的中心矩为

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}}}$$

其中, $\gamma = \frac{p+q}{2} + 1, p+q=2, 3, \dots$ 。归一化的中心矩具有平移和伸缩不变性。组合归一化的二阶和三阶中心矩可得到以下 7 个不变矩 $\Phi_1 \sim \Phi_7$:

$$\begin{aligned} \Phi_1 &= \eta_{20} + \eta_{02} \\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \lfloor (\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \rfloor + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \lfloor 3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \rfloor \\ \Phi_6 &= (\eta_{20} - \eta_{02}) \lfloor (\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \rfloor + 4\eta_{11}^2 (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \lfloor (\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \rfloor + \\ &\quad (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \lfloor 3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \rfloor \end{aligned}$$

由于不变矩 $\Phi_1 \sim \Phi_7$ 对图像的平移、伸缩及旋转均不变, 因此代表了模式形状的基本特征, 可将目标图像的上述 7 个不变矩的测量值作为其特征矢量。在具体的模式识别过程中, 也可根据所识别图像客体的具体性质选择不同的不变矩的组合, 或者将不变矩与其它的特征量相结合。

3 ART-2 神经网络

3.1 ART 神经网络简介

ART 网络是基于美国 Boston 大学的 S.Grossberg 和 A.Carpenter 于 1976 年提出的自适应共振理论(Adaptive Resonance Theory)构造的, 其目的就是为了解决用人工神经网络模拟人脑的自组织记忆与稳定性的问题。它利用生物神经细胞之间自兴奋和侧抑制的动力学原理, 让输入模式通过网络双向连接权进行识别与比较, 最后达到共振来实现记忆, 并以相同方法实现网络的联想。ART 网络是以认知和行为模式为基础的无教师、矢量聚类、竞争学习型网络, 它成功地解决了网络学习稳定性(固定某一分类集)和可塑性(调整网络固有参数的学习状态)之间的矛盾。当每个输入模式提供给网络时, 将其与该模式最接近的原型模式相比较, 如果它们不足以匹配, 那么它将作为一个新的原型模式而被选中存储。自提出自适应共振理论以来, 先后构造了 ART-1、ART-2、ART-3、Fuzzy ART 等多种神经网络模型。ART-1 仅适用于二进制的输入样本模式, ART-2 可适用于任意的模拟输入信号模式。

3.2 ART-2 神经网络结构

ART2 网络结构如图 1 所示, 整个网络分为注意子系统和取向子系统两部分。注意子系统包括比较层 F_1 和识别层 F_2 以及 F_1 与 F_2 之间的连接通路—自适应长期记忆(LTM)。取向子系统则由重置机构组成。比较层包含 6 个子层(分别是 w, x, v, u, p, q 层)以及 3 个增益控制模块(G_1, G_2, G_3)。所有的子层, 以

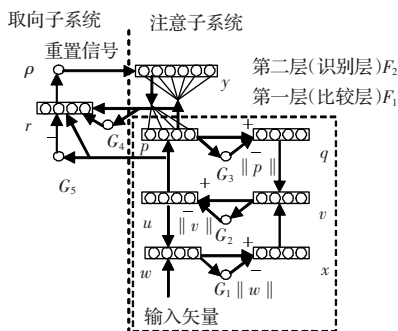


图 1 ART-2 神经网络结构

及取向子系统中的 r 层所包含的神经元数目均相同。 F_1 中的所有子层之间的联接方式为对应神经元之间的一对一联接。

3.3 ART-2 神经网络的基本原理

设 I_i 为输入特征向量的第 i 个分量, 每个输入分量在 w 中对应一个输入神经元 w_i , 其输入方程为

$$w_i = I_i + a u_i$$

其中 a 为正反馈系数。 F_1 层有两个正反馈回路, 其目的是抑制噪声并增强有用信号。

w, x, v, u 构成第一个正反馈回路, 完成两次规格化运算和一次非线性运算。对输入矢量的规格化运算为

$$x_i = \frac{w_i}{e + \|w\|}$$

其中, e 为一很小的正数, 其作用是当没有输入信号时使活化能保持有限。 v_i 神经元的非线性运算为

$$v_i = f(x_i) + b f(q_i)$$

非线性函数 $f(x)$ 决定了 F_1 层的对比度增强特性, 同时增强系统的抗干扰能力, 通常采用阈值函数

$$f(x) = \begin{cases} 0 & 0 \leq x \leq \theta \\ x & x > \theta \end{cases} \quad (0 < \theta < 1)$$

θ 为滤波因子, 对噪声抑制能力起决定作用, θ 取值越高, 抑制能力越强, 但也削弱了有用信号, 一般 θ 取值 $0 < \theta \leq 1/\sqrt{M}$, M 为特征矢量的维数。 v_i 的输出经 u 层神经元 u_i 进行规格化运算

$$u_i = \frac{v_i}{e + \|v\|}$$

u, p, q, v 构成第二个正反馈回路, p 层神经元向 F_2 输出激励, 同时也接收 F_2 自上而下的输入, 其方程为

$$p_i = u_i + \sum_j g(y_j) w_{ij}$$

$g(y)$ 为 F_2 层的输出函数, w_{ij} 为自上而下的 LIM 系数。 q 层神经元 q_i 实现对 p_i 的规格化

$$q_i = \frac{p_i}{e + \|p\|}$$

识别层为一个相互竞争网络。从下到上的输入计算式为:

$$T_j = \sum_i p_i w_{ji}$$

F_2 层的输出函数的形式如下:

$$g(y_j) = \begin{cases} d & T_j = \max\{T_k\} \quad \forall k \\ 0 & \text{其它} \end{cases}$$

这里 F_2 层采用的是独活的竞争机制, 即只有输入向量和联接权的点积最大的那个神经元才会产生输出。因此 P 层的活化能方程式可以改写为:

$$p_i = \begin{cases} u_i & F_2 \text{ 未激活} \\ u_i + d x_{ij} & F_2 \text{ 上第 } j \text{ 个神经元激活} \end{cases}$$

由此可见, 输入矢量经 F_1 层两个正反馈回路的正反馈滤波作用, 噪声被抑制, 有用信号被增强, 并对 F_2 层产生自下而上的输出, 而 F_2 受 F_1 的激励, 经独活的竞争之后, 仅一个神经元获胜, 并将它的信息 w_{ij} 返回 F_1 , 在 p 层与输入模式进行比较, 由取向子系统根据比较后的相似度判断其输入模式是否为与获胜神经元相对应的特征模式。

取向子系统 r 层的活化能方程式为

$$r_i = \frac{u_i + c p_i}{\|u\| + \|c p\|}$$

则相似度定义为

表1 不变矩数据及识别结果

图形	不变矩特征 $\Phi_1 \sim \Phi_7$	识别结果	相似度
接头	0.159 851 618 310 54 0.000 050 252 948 86 0.000 000 588 897 09	00	0.999 999 964 703 83
螺杆	0.217 614 127 312 29 0.011 062 726 711 71 0.004 892 150 418 24	01	0.999 947 523 676 03
螺母	0.411 103 130 334 02 0.000 213 666 051 13 0.000 001 099 550 88	10	0.999 981 261 115 32
垫片	0.167 198 917 937 47 0.000 016 312 576 43 0.000 000 910 248 59	11	0.999 999 964 226 54

$$\|r\| = \frac{[1+2\|cp\|\cos(u,p)+\|cp\|^2]^{1/2}}{1+\|cp\|}$$

如果 $\|r\| < \rho$ (ρ 为警戒参数且 $0 < \rho \leq 1$), 则 p 与 u 失配, 取向子系统会向第二层发出重置信号, 同时使前一个获胜神经元长时期抑制, 并产生新一轮竞争。如果 $\|r\| \geq \rho$, 即为达到共振, 则输入模式即为获胜神经元所代表的特征模式, 此时将 LTM 系数按下式学习

$$\text{从下到上: } \dot{w}_{ji} = g(y_j)(p_i - w_{ji})$$

$$\text{从上到下: } \dot{w}_{ij} = d(u_i + dw_{ij} - w_{ij})$$

在快速学习模式下, 当网络达到平衡即 $\dot{w}_{ji} = \dot{w}_{ij} = 0$ 时, 有

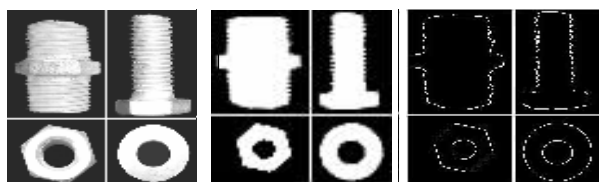
$$w_{ij} = w_{ji} = \frac{u_i}{1-d}$$

若输入模式与学习过的所有模式都失配, 将在 F_2 层开辟一个新的输出端, 并保存其 LTM 系数。

4 实验和结果

4.1 图像采集及预处理

本实验采用 CCD 摄像机及 DH-CG410 图像采集卡。在图像处理中对图像进行识别或匹配, 一般先计算直方图, 确定阈值, 然后对图像进行二值化处理, 对处理后的图像进行不变矩的计算。在二值化的图像计算时, 直接利用每个位点像素来判断是否为图像点, 如果像素不是目标物体图像点则不必进行计算。如图 2 是对 4 种零件图像进行滤波、二值化、边缘检测等一系列的变换得到的结果图像。



(a) 零件原图像 (b) 图像二值化结果 (c) 图像边缘检测结果

图2 零件图像处理结果

4.2 图像特征提取

分别以上述 4 种模式图像以及它们旋转 $30^\circ, 45^\circ, 60^\circ, 90^\circ, 150^\circ$ 以及扩大 1.2 倍, 1.5 倍, 2 倍后的 36 幅图像作为训练样本, 计算各图像的 7 个不变矩 $\Phi_1 \sim \Phi_7$ 构成其特征向量, 在此采用 7 个不变矩中 $\Phi_1 \sim \Phi_3$ 的作为图像特征, 计算结果如表 1 所示。

4.3 基于 ART-2 神经网络的图像模式识别

首先将所有神经元的输出设为 0 向量, 对参数 a, b, c, d 进行初始化, $a, b > 0; 1 \geq d \geq 0; \frac{cd}{1-d} \leq 1$ (比率越接近 1, 网络对误

配越敏感, 因为在其他条件不变的情况下, $\|r\|$ 越小); 阈值 $0 < \theta \leq 1/\sqrt{M}$; 警戒参数 $0 < \rho < 1; e < 1$ 。经多次调节, 选取 $a=8; b=8; c=0.12; d=0.9; \theta=0.12; \rho=0.995$ 。将 F_2 层至 F_1 层权值矩阵 w_{ij} 初始化为零矩阵, 而 F_1 层至 F_2 层权值矩阵 w_{ji} 按照公式 $w_{ji}(0) \leq \frac{1}{(1-d)\sqrt{M}}$ 进行初始化。输入向量采用图像 7 个不变

矩中的前 3 个作为训练样本 (通常可选 3 个或 4 个不变矩)。将第一种模式 (接头) 的不变矩输入到 ART2 网络的第一层进行预处理, F_1 层稳定后, 将数据输入到 F_2 识别层, 由于 F_2 层未被激活, 所以自动将这种模式存入第一种模式 (即 w_{ij} 的第一列); 将第二种模式 (螺杆) 的不变矩输入到 ART2 网络, 第一层进行预处理后, 如果 $\rho/(e + \|r\|) > 1$, 即不满足当前的警戒参数, 则向 F_2 层送出一个重置信号, 把当前激活的 F_2 层神经元排除, F_2 层开辟一个新的模式存储当前模式 (即 w_{ij} 的第二列), 同理可训练第 3 种模式 (螺母) 和第 4 种模式 (垫片) 以及在实际应用中的新模式。

实验采用了 4 种零件的图像样本, 对上述神经网络进行识别检验, 识别的正确率达到了 100%。表 1 为零件原图像进行识别时得到的检测数据。实验中还对零件图像进行旋转若干角度且扩大若干倍后, 正确率仍能达到 100%。

5 总结

本文针对 BP 神经网络在应用于模式识别时, 面对新输入模式时, 必须重新训练网络的缺点, 提出了应用 ART2 神经网络进行模式识别的方法。实验证明 ART-2 本身具有很强的自适应学习能力, 可以存储记忆新模式而不需重新训练整个网络, 克服了 BP 神经网络的上述缺点, 同时取得了很好的识别结果。这在机器人视觉伺服控制中具有很高的实用价值。

(收稿日期: 2007 年 8 月)

参考文献:

- [1] Zitova B. Robust detection of significant points in multi-frame image[J]. Pattern Recognition Letters, 1999, 20: 199-206.
- [2] Sherstinsky A, Picard R W. On the efficiency of the orthogonal least square straining method for radial basis function networks[J]. IEEE Trans Neural Networks, 1996, 7(1): 195-200.
- [3] 王英健, 戎丽霞. 基于遗传 BP 算法的神经网络及其在模式识别中的应用[J]. 长沙交通学院学报, 2005, 25(1): 53-56.
- [4] Ao G, Akazawa H, Izumi M, et al. A method of model-based object recognition[C]. Japan/USA Symposium on Flexible Automation, 1996, 2: 905-912.
- [5] 范立南, 徐心和. 基于不变矩特征和神经网络的图像模式模糊分类[J]. 东北大学学报, 2004, 25(8): 738-741.