

基于流水线图像处理结构的多层模板相关神经网络*

安向京 常文森

(国防科技大学自动控制系 长沙 410073)

摘要 在自主地面车辆中,视觉系统的重要作用之一是根据路标来定位。本文提出了一种便于流水线图像处理结构实现的多层模板相关神经网络(MTCNN)。文中给出了MTCNN的基本结构及训练算法,并且将其与经典的多层前馈神经网络(MLFNN)进行了比较。仿真结果表明,本文提出的算法结构在多层前馈神经网络的分类能力与采用通用图像处理硬件的可实现性之间,取得了良好的折衷。

关键词 模板匹配,神经网络,流水线图像处理器,路标

分类号 TP391

Multi-layer Template Correlation Neural Network Based on Pipelined Image Processing Structure

An Xiangjing, Chang Wensen.

(Department of automatic control, NUDT, Changsha, 410073)

Abstract It is one of the important tasks of the vision system of an autonomous land vehicle (ALV) to locate itself by lane mark. In this paper, a multi-layer template correlation neural network (MTCNN) based on the pipelined image processing structure is proposed for the recognition of lane mark. A structure of the MTCNN and the training algorithm are presented. In addition, a comparison between MTCNN and MLFNN is introduced. The results of simulation manifest that the proposed MTCNN is very efficient for the task such as recognition of lane mark that is based on the pipelined image processing structure.

Key words template match, neural network, pipelined image processor, lane mark.

在自主地面车辆的研究中,视觉系统的重要作用之一是根据路标来定位^[1]。通常,自主地面车辆内部有用于导航和定位的地理信息系统。路标是一种对特殊道路结构的描述,它不仅仅是出现在特定道路结构前的路面上的人工标志,而且在地理信息系统中也有对应的可标识并被查询的特征数据结构。利用被识别了的路标以及它和自主地面车辆之间的相对定位关系,就可以确定自主地面车辆在地理信息系统中的位置了。应用于自主地面车辆中的视觉系统,必须满足大吞吐量和强实时性的要求,这便对视觉系统的硬件提出了很高的要求。使用流水线图像处理器被认为是能够满足这些要求的有效办法,尤其是在低层图像处理中,使用流水线图像处理器来解决实时图像处理问题已经很普遍了^[2]。

我们实现的路标识别定位系统,本质上是以道路景物和路标图像为先验知识的实时图像处理和理解系统,其核心问题是模式识别中的分类问题。近来,有很多关于神经网络分类器和统计分类器的比较研究^[3,4]。在统计分类器中,模板匹配是一个重要方法。采用模板匹配方法,通常需要将输入样本逐一与标准模板相比较,从中选择出匹配程度最大的。但它是基于相关运算的,相关运算的明显缺点是运算量太大,并且,模板匹配中,模板的尺寸通常不能太大。采用神经网络的方法,为避免过多的输入结点带来的复杂的稳定性和收敛性问题,通常不是将整个图像作为输入,而是对输入图像进行一些预处理,提取图像的结构特征作为网络的输入。然而,并不是所有的算法都是易于用流水线图像处理器实现的,我们的目标就是充分利用流水线图像处理器的硬件特点,设计出实时高效的路标识别算法。

模板匹配算法的效率可以通过以下途径提高。首先,采用二分法的思路,模板可以根据输入模式的

* 国家部委基金项目资助

1998年12月25日收稿

第一作者:安向京,男,1973年出生,博士生

主要特征来设计,从而使标准模板集可以逐次被等分成两部分。如果将模板匹配运算与只有两层并且神经活化函数为线性函数的多层前馈神经网络作比较,可以发现模板系数与神经网络的权系数是等价的。其次,受到叠代相关^[5]思路的启发,可以不是从一次相关的结果中得到输入模式的类型,而是将这一次相关的结果作为输入再次做模板相关,从后续的第二次,第三次等的结果中得到输入模式的类型。这样一来,模板尺寸受限的问题就能得到一定程度地解决了。最后,将非线性特性引入到模板相关的过程中。在每次模板相关之后,定义了神经活化函数,将经过神经活化函数映射之后的结果作为下一次模板相关的输入。综合上述讨论,我们提出了一种具有很多神经网络特性的模板相关算法,称之为多层模板相关神经网络。

1 多层模板相关神经网络

1.1 基本结构

多层模板相关神经网络是一种兼具模板匹配和多层前馈神经网络的部分特点的运算结构。它对输入模式依次做多次不同的模板相关,并且对每次模板相关的中间结果进行神经活化函数映射,最终得到的结果值即反映了输入模式的某种分类特征。

在这里,我们首先定义模板相关算子^[6]为:

$$(w \circ A)(i, j) = \sum_{m=0}^{M-1-N-1} \sum_{n=0}^{N-1-1} w(m, n) A(i+m, j+n) \tag{1.1}$$

其中: $w(i, j)$ 为相关模板, $i=0, 1, \dots, M-1, j=0, 1, \dots, N-1$. $A(i, j)$ 为被相关的图像, $i=0, 1, \dots, I-1, j=0, 1, \dots, J-1$. $(w \circ A)(i, j)$ 为模板相关的结果, $i=0, 1, \dots, I-M, j=0, 1, \dots, J-N$.

我们提出的多层模板相关神经网络可以表示成递推定义的形式如下,此处只考虑各层模板尺寸相同的情形。

$$\begin{aligned} A_l(i, j) &= f_l(w_{l-1}(i, j) \circ A_{l-1}(i, j)) \\ A_1(i, j) &= f_1(w_0(i, j) \circ A_0(i, j)) \end{aligned} \tag{1.2}$$

其中: l 是第 l 层模板相关神经网络的层序号, $l=0, 1, \dots, L$. f_l 是第 l 层模板相关神经网络的神经活化函数。 $w_{l-1}(i, j)$ 为第 $l-1$ 层的相关模板, $i=0, 1, \dots, M-1, j=0, 1, \dots, N-1$, $A_{l-1}(i, j)$, 为第 $l-1$ 层的输入图像, $i=0, 1, \dots, I-(l-1)(M-1)-1, j=0, 1, \dots, J-(l-1)(N-1)-1$. $A_l(i, j)$ 是第 l 层模板相关神经网络的中间结果, $i=0, 1, \dots, I-l(M-1)-1, j=0, 1, \dots, J-l(N-1)-1$. 上面, L 为该模板相关神经网络的总层数。为了叙述的方便,我们将第 l 层模板相关神经网络输入的定义域记为:

$$Def_l = \{(i=0, 1, \dots, I-l(M-1)-1, j=0, 1, \dots, J-l(N-1)-1)\} \tag{1.3}$$

上面 (1.2) 式可以表示成为另一种形式:

$$A_l(i, j) = f_l(w_{l-1}(i, j) \circ f_{l-1}(w_{l-2}(i, j) \circ \dots \circ f_1(w_0(i, j) \circ A_0(i, j)) \dots)) \tag{1.4}$$

这些表达式所表达的运算关系可以用图1清楚地表示出来。

1.2 训练方法

设有 L 层多层模板相关神经网络,并且整个网络的输出为单像素值: $A_L(i, j), i=0, j=0$. 整个多层模板相关神经网络使用相同尺寸的不同模板 $w^l(m, n), m=0, 1, \dots, M-1, n=0, 1, \dots, N-1$.

给定训练对集:

$$Set = \{(A_0(i, j)_{k, y_k})\}, k = 1, 2, \dots, K \tag{1.5}$$

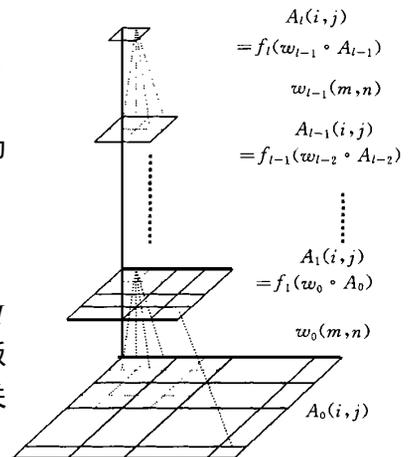


图1 多层模板相关神经网络示意图
Fig 1 Demonstration of MTCNN

其中: $A_0(i, j)_k$ 为第 k 对训练对中的输入模式, $(i, j) \in Def^0, y^k$ 为网络的期望输出。

使用上述训练对集, 对某一输入 $A_0(i, j)_k$, 网络的期望输出为 y^k , 网络的实际输出为 $A_L(i, j)_k$ 。模板参数的选择要使得下面的误差方程取得最小值。

$$E = \sum_{k=1}^K E_k \quad (1.6)$$

$$E_k = e(A_L(i, j)_k, y^k) \quad (1.7)$$

显然, E 是关于多层模板相关神经网络中所有模板参数的函数, 现在要寻找使 E 最小的模板集:

$Template = \{W_0, W_1, \dots, W_{L-1}\}$ 。这是一个非线性优化问题, 考虑使用最速下降法来进行逐层模板学习寻优, 即:

$$w_l(m, n)_{t+1} = w_l(m, n)_t - \mu \sum_{k=1}^K \frac{\partial E_k}{\partial w_l(m, n)_t} \quad (1.8)$$

其中, μ 是学习速率, $w_l(m, n)_t$ 为第 t 次学习时第 $l+1$ 层的模板。现在的关键问题是求 $\frac{\partial E_k}{\partial w_l(m, n)_t}$, 我们可以得到下式:

$$\frac{\partial E_k}{\partial w_{L-1}(m, n)} = (e_{A_L(i, j)} \cdot (f_{L-1}^{w_{L-1}(i, j)} \cdot A_{L-1}(i+m, j+n)))$$

以下在不至于混淆的情况下, 作如下简化记法:

$$e = e_{A_L(i, j)}$$

$$f^l = f_{w_{l-1}(i, j)}^{A_{l-1}(i, j)}, \quad l = 1, 2, \dots, L$$

也就是有:

$$\frac{\partial E_k}{\partial w_{L-1}(m, n)} = e \cdot f^L \cdot A_{L-1}(i+m, j+n)$$

$$\frac{\partial E_k}{\partial w_{L-2}(m, n)} = e \cdot f^L \cdot (w_{L-1} \cdot (f^{L-1} \cdot A_{L-2}(i+m, j+n)))$$

$$\frac{\partial E_k}{\partial w_{L-3}(m, n)} = e \cdot f^L \cdot (w_{L-1} \cdot (f^{L-1} \cdot (w_{L-2} \cdot (f^{L-2} \cdot (A_{L-3}(i+m, j+n))))))$$

$$\frac{\partial E_k}{\partial w_0(m, n)} = e \cdot f^L \cdot (w_{L-1} \cdot (f^{L-1} \cdot (w_{L-2} \cdot \dots \cdot (f^1 \cdot (A_0(i+m, j+n)))))) \quad (1.9)$$

注意上式中, $A_l(i+m, j+n)$ 中 $(i, j) \in \{(i, j) \mid (i+M, j+N) \in Def^l\}$

上述计算过程是最速梯度下降法的寻优过程, 有很多文章讨论了该算法在收敛速度等方面的不足^[7], 在此我们采用自适应学习速率调节的方法, 来改进多层模板相关神经网络模板参数的学习效率。为此构造模板参数及学习速率的调节规则为:

$$w_l(m, n)_{t+1} = \begin{cases} w_l(m, n)_t & E_n \geq E_{n-1} \\ w_l(m, n)_t - \mu_t \sum_{k=1}^K \frac{\partial E_k}{\partial w_l(m, n)_t} & E_n < E_{n-1} \end{cases} \quad (1.10)$$

$$\mu_{t+1} = \begin{cases} \mu_t \cdot dm & E_n \geq E_{n-1} \\ \mu_t \cdot im & E_n < E_{n-1} \end{cases}$$

其中: dm 为学习速率的增加速率, $dm \in (0, 1)$, im 为学习速率的减小速率, $im \in (1, +\infty)$, μ_t 为第 t 次学习时的学习速率。

2 仿真性能测试

为验证多层模板相关神经网络进行模式分类思路的可行性, 以及上述学习速率自适应调整的模板训练算法的有效性, 将上述 MTCNN 与经典的多层前馈神经网络 (Multi-Layered Feedforward Neural Network) 进行了对比。

根据前文的叙述, 给定待识别的标准路标为图2所示四种。这里, 定义 MTCNN 的结构如前文所述,

设 $L = 2$, 即经过两次连续的模板相关后, 整个网络的输出为单像素值。 $M = N = 8$, 即两层模板的尺寸均为 8×8 。 $I = J = 15$, 即原始输入模式的尺寸为 15×15 。MLFNN 的结构参见文 [9], 在这里, 我们定义的 MLFNN 含有一个中间隐层, 中间隐层包含三个中间单元, 不含偏置项, 输入为 15×15 共 225 个单元, 输出为 $G = 4$ 个单元。

两种网络采用相同的误差函数及神经元活化函数:

$$\begin{cases} e_k(\hat{y}_k, y_k) = (\hat{y}_k - y_k)^2 \\ f(x) = x \end{cases} \quad (2.1)$$

其中, \hat{y}_k 对 MTCNN 为 $A_2(0, 0)_k$, 对 MLFNN 为网络实际的输出向量。 y_k 对 MTCNN 为第 k 个教师值, 对 MLFNN 为第 k 个教师向量。

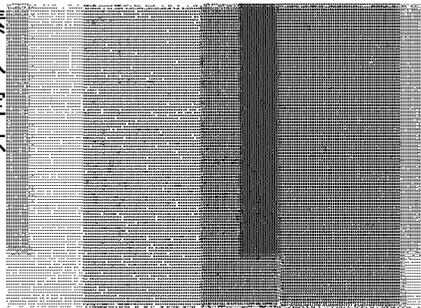


图2 标准路标

Fig 2 Normalized lane marks

对于 MTCNN, 给定对应于图2中的各个路标的教师值分别为 $(0.15, 0.35, 0.55, 0.75)$ 形成训练对集。对于 MLFNN, 指定第 i 个路标的教师值为: 网络输出的第 i 个分量为 1 其余为 0 的向量形成

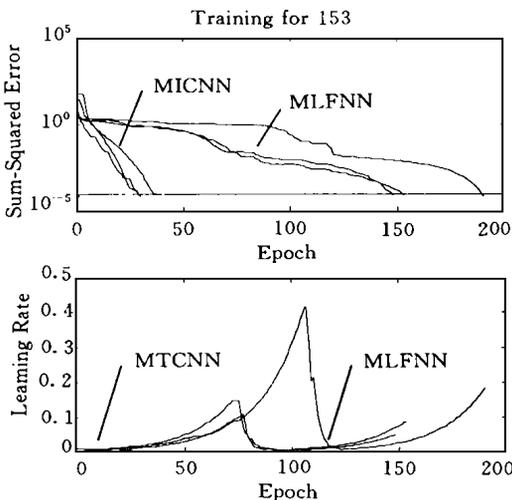


图3 训练过程对比

Fig 3 Comparison of training procedures

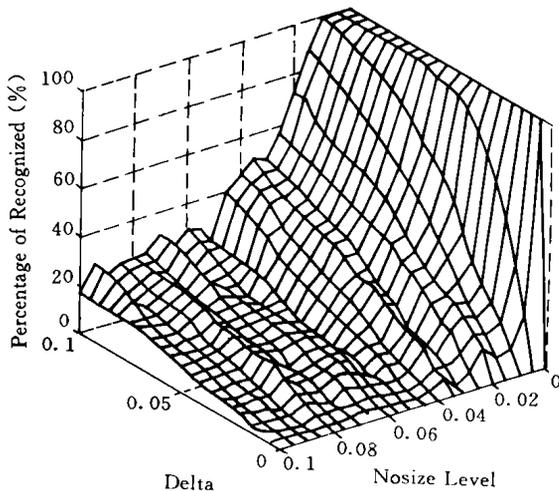


图4 识别率变化趋势

Fig 4 Trend of percentage of recognized

的训练对集。

对于 MTCNN 我们实现了前文所述的学习算法, 对 MLFNN 采用了文 [8] 所述算法。我们使用上述的标准路标模式训练对, 对 MTCNN 及 MLFNN 分别进行了训练, 在相同的初始控制参数下, 训练算法的误差变化曲线和学习速率变化曲线对比如图3所示。

为了增加识别网络的鲁棒性, 对原始输入图像叠加了标准正态分布的噪声干扰, 叠加方法为:

$$A_0(i, j)_k = A_0(i, j)_k + Rand(i, j)_k \cdot nl$$

其中, $rand(i, j)_k$ 是与 $A_0(i, j)_k$ 同尺寸的矩阵, 其元素为标准正态独立同分布, nl 为噪声水平。用新的训练对集对 MTCNN 及 MLFNN 分别进行了训练, 在相同的初始控制参数下, 训练算法的误差变化曲线和学习速率变化曲线的趋势与图3类似, 只是训练次数显著增加。

值得注意的是, 由于两种网络结构的输出形式不同, 识别结果的统计方法也不同。对 MTCNN, 我们认为当输出结果值落在给定的期望教师值附近的一个邻域 $(-\delta, \delta)$ 内, 便被确认为是该模式, 显然, $\delta = (0, \min(\frac{y_i - y_j}{2}))$, $i, j = 1, 2, \dots, G$ 。对 MLFNN, 我们认为输出向量中值最大的分量所对应的序号, 就被确认为是序号所对应的模式。如图4, 我们得到了当 $\delta = (0, 0.1)$ 时, MTCNN 对叠加了不同水平的噪声的输入模式的正确识别率的变化趋势。从图中, 可以看出 δ 取 $\delta = (0.01,$

0.04), 即可得到较满意的识别率和分类能力。

3 结论与展望

本文提出了一种便于流水线图像处理器实现的模式分类运算结构——多层模板相关神经网络。它提供了在通用图像处理器上应用神经网络的一种新思路, 这使它不仅具有通用处理器上的算法的灵活性和廉价性, 而且具有与专用处理器上的算法相当的实时性和处理能力, 在与模式识别相关的应用领域有着很大的潜在应用前景。

参考文献

- 1 Devy M, Chatila R, Fillatreau P, Lacroix S, Nashashibi F. On autonomous navigation in a natural environment. *Robotics and Autonomous Systems*, vol. 16, no. 1, pp. 5- 16, 1995
- 2 Olson T J, Taylor J R. Programming a pipelined image processor. *Computer Vision and Image Understanding*, vol. 64, no. 3, November, pp. 351- 367, 1996
- 3 Ripley B. Neural network and related methods for classification. *J. Roy Statist. Soc., Series B*, vol. 56, no. 3, pp. 409- 456, 1994
- 4 Idan Y, Auger J M, Darbel N, Sales M, Chevallier R, Dorizzi B, and Cazuguel G. Comparative study of neural networks and non-parametric statistical methods for off- line hand- written character recognition. In *Proc. Int. Conf. Artificial Neural Networks, I. Aleksander and J. Taylor, Eds. Brighton, U. K.: North- Holland*, vol. 2, Sept. 1992, pp. 1607- 1610
- 5 Burt P J. Smart sensing within a pyramid vision machine. *Proc. IEEE*, vol. 76, no. 8, pp. 1006- 1015, 1998
- 6 Chou C H, Chen Y C. A VLSI architecture for real- time and flexible image template matching. *IEEE Trans. Circuits and Systems*, vol. 36, no. 10, October, pp. 1336- 1342, 1989
- 7 梁民, 孙仲康. 多层前馈神经网络的快速学习算法及其仿真研究. *系统工程与电子技术*, 1993, 9: 47
- 8 史忠植. *神经计算*. 北京: 电子工业出版社, 1993