

文章编号:1007-130X(2012)02-0116-07

# 基于 Hermite 神经网络的动态手势学习和识别<sup>\*</sup>

## Dynamic Gesture Learning and Recognition Based on the Hermite Neural Network

李文生<sup>1</sup>, 解梅<sup>1,2</sup>, 邓春健<sup>1</sup>, 姚琼<sup>1</sup>

LI Wen-sheng<sup>1</sup>, XIE Mei<sup>1,2</sup>, DENG Chun-jian<sup>1</sup>, YAO Qiong<sup>1</sup>

(1. 电子科技大学中山学院计算机工程系, 广东 中山 528402;

2. 电子科技大学电子工程学院, 四川 成都 610054)

(1. Department of Computer Engineering, Zhongshan Institute,

University of Electronics Science and Technology of China, Zhongshan 528402;

2. School of Electronic Engineering, University of Electronics Science

and Technology of China, Chengdu 610054, China)

**摘要:**为提高动态手势学习速度和识别准确率,本文提出一种基于 Hermite 正交基前向神经网络的动态手势识别方法。利用 Camshift 算法实时跟踪手势运动轨迹,提取手势特征向量作为神经网络的输入;以 Hermite 正交基函数作为隐含层激励函数构造三层前向神经网络,并给出一种基于伪逆的直接计算权值方法和根据网络目标精度要求自适应确定隐含节点数目方法;运用训练好的 Hermite 神经网络识别动态手势。测试结果表明:Hermite 神经网络能够提高网络的学习训练速度和精度,提高手势学习速度和识别准确率,而且在手势识别方面具有较好的鲁棒性和泛化能力。

**Abstract:** In order to improve the training speed and identification accuracy, a method of dynamic gesture recognition based on the Hermite orthogonal basis feed-forward neural network is put forward. At first, the CamShift algorithm is used to track trajectories of moving fingertips and the characteristic vector of gesture is extracted as the input of the neural network. Then, a feed-forward neural network which hides the layer neurons is activated by a group of Hermite orthogonal polynomial functions is, and a method to determine the network weights directly and determine the number of hidden layer nodes adaptively is proposed. Finally, gestures based on machine vision are recognized through the trained Hermite neural network. The experimental results show that the Hermite neural network can enhance the speed and precision of network training, improve the learning speed and identification accuracy of gesture recognition and has good robustness and generalization ability.

**关键词:** Hermite 神经网络;权值直接确定;隐含结点数自适应确定;指尖跟踪;动态手势识别

**Key words:** Hermite neural network; weights' direct determination; hidden node number adaptive determination; fingertips tracking; dynamic gesture recognition

doi:10.3969/j.issn.1007-130X.2012.02.022

中图分类号:TP391

文献标识码:A

\* 收稿日期:2010-02-18;修订日期:2011-04-28

基金项目:广东省自然科学基金资助项目(8152840301000009);广东省科技计划项目(2009B030803031)

通讯地址:528402 广东省中山市电子科技大学中山学院计算机工程系

Address: Department of Computer Engineering, Zhongshan Institute, University of Electronics Science and Technology of China, Zhongshan, Guangdong 528402, P. R. China

## 1 引言

BP 多层前向神经网络对于多输入、多输出非线性映射关系问题具有强有力的刻画和建模能力,因为它具有高度的并行性、分布式存储、自适应学习能力、便于用超大规模集成电路实现等特点,同时具有较强的泛化能力和容错能力,所以广泛应用于模式识别、信号处理、智能控制等领域<sup>[1,2]</sup>。

动态手势识别是当前模式识别和人机交互领域中的热点研究课题<sup>[3~9]</sup>。目前,动态手势识别多以 BP 神经网络作为识别方法<sup>[7~9]</sup>。但是,BP 神经网络算法存在一些固有的缺点,如容易陷入局部极小、收敛速度慢、隐含层节点数目确定缺乏理论指导等<sup>[1,2]</sup>。许多改进算法不断被提出来,它们大体上可以分为两类<sup>[1]</sup>:

(1)基于标准梯度下降法的改进(如共轭梯度法、带动量的梯度下降法等);

(2)基于数值优化方法的改进(如拟牛顿法、LM 方法等)。

这些算法大都把注意力放在网络训练迭代规则的改进上,尽管有所改善,但不能完全克服学习训练中的冗长迭代过程,不太适应需要对不同用户手势或新的手势进行在线学习的情况。

文献[10]提出了基于正交基多项式函数的神经网络构造理论和构造方法。文献[11]用构造性方法证明了对于任意的  $n$  阶多元多项式函数,可以构造一个三层前向神经网络以任意精度逼近该多项式,而且所构造网络的隐含层节点数目仅与多项式的维数和阶数有关,并且提出了通过改进隐含层神经元的激励函数和网络定义来解决前向神经网络的局部极小和收敛速度慢问题的思想和方法。文献[12]在此基础上,提出了单输入单输出(SISO)情况下的基于 Hermite 正交基的前向神经网络,并且针对该网络模型提出一种基于伪逆的直接计算权值方法以及依照训练精度要求自适应确定隐含层节点算法,大大提高了网络的学习训练速度,但可惜的是它适合一元函数的逼近,但不适应多输入、多输出(MIMO)的手势识别情况。

本文讨论多输入、多输出的 Hermit 正交基前向神经网络,并将其应用于基于视觉的手势识别。仿真实验结果表明,本文提出的方法在动态手势识别中训练学习速度快、识别正确率高,并且具有较好的鲁棒性和泛化能力,效果明显优于传统的 BP 神经网络。

## 2 手势样本的获取和表示

要实现基于机器视觉的手势识别,必需首先实现多个指尖目标的跟踪,并由此得到手势的运动轨迹。Bradski 提出的 CamShift(Continuously adaptive mean shift,简称 CamShift)算法是一种以颜色直方图为基础的目标跟踪算法,它能自动调节搜索窗口大小以适应被跟踪目标的大小,并可以有效地解决目标变形和遮挡的问题<sup>[13]</sup>。本项目采用 CamShift 算法来对戴上特定颜色指套的多个指尖目标进行检测和跟踪。

### 2.1 指尖颜色的学习

在进行指尖目标检测和跟踪之前,需要对指尖的颜色进行学习训练。在涉及颜色处理时,常用的颜色空间有 RGB、HSV 等。RGB 颜色空间中红绿蓝三原色受光照强度影响较大,三原色之间存在相关性。同时,在该颜色空间下,两点间的欧氏距离与颜色实际距离不成线性比例,不能很好地衡量颜色间的差异。与 RGB 颜色空间相比,HSV 颜色空间除亮度(V)属性外,其他两个基本属性色相(H)和饱和度(S)受光照影响较小。为了降低光照影响,同时降低具有相似色相背景的影响,使用 HSV 颜色空间的色相(H)和饱和度(S)来对指尖目标颜色进行学习训练,得到一个二维的颜色直方图(二维查找表)。

首先用户将指尖放在一个小红色矩形框指定的地方,并让该红色矩形框停留在指尖上一段时间,然后通过遍历该矩形框中所有像素,统计各种颜色出现的次数,并进行归一化处理得到颜色直方图,该颜色直方图直观地表达了指尖目标在色彩空间的分布状况。当然,为了提高系统对光照变化、指尖形状和位置变化的鲁棒性,我们可以在不同的光照条件下,对不同的指尖进行颜色训练学习。

### 2.2 指尖目标的检测

首先选择初始搜索窗口,使窗口包含整个跟踪目标,对搜索窗口中的每一个像素,通过查询二维颜色直方图,可以得到该像素为目标像素的概率。经上述处理,视频图像转换为目标颜色概率分布图  $I(x, y)$ ,这个过程称为反投影运算。

为简化后续跟踪计算,可以通过设置适当的阈值将目标颜色概率分布图  $I(x, y)$  二值化,得到相应的二值图像  $B(x, y)$ 。

$$B(x, y) = \begin{cases} 0, & I(x, y) < p \\ 1, & I(x, y) \geq p \end{cases} \quad (1)$$

其中  $p$  是给定的阈值。

显然,二值图像  $B(x, y)$  给出了需要关注的指尖目标像素。

### 2.3 指尖目标的跟踪

通过 CamShift 算法来对具有特定颜色的指尖目标跟踪定位。CamShift 算法是一种运动目标跟踪算法,它是建立在 MeanShift 算法的基础上<sup>[13]</sup>。

Mean Shift 算法可以分为以下 4 步:

(1) 选择搜索窗的大小和初始位置,设  $(x_0, y_0)$  为初始搜索窗口中心。

(2) 在搜索窗口中计算二值图像  $B(x, y)$  的零阶矩  $M_{00}$ ,一阶矩  $M_{10}$ 、 $M_{01}$  和质心  $(x_c, y_c)$ 。

$$\begin{aligned} M_{00} &= \sum_x \sum_y B(x, y), \\ M_{10} &= \sum_x \sum_y xB(x, y), \\ M_{01} &= \sum_x \sum_y yB(x, y) \end{aligned} \quad (2)$$

$$x_c = M_{10}/M_{00}, y_c = M_{01}/M_{00} \quad (3)$$

(3) 调整搜索窗口的中心到质心位置,即:  $x_0 = x_c, y_0 = y_c$ 。

(4) 重复步骤(2)和(3),直到搜索窗口中心会聚,即每次窗口移动的距离小于一定的阈值:  $|x_0 - x_c| < \epsilon$  并且  $|y_0 - y_c| < \epsilon$ 。

MeanShift 算法的缺点是不能自适应地调整搜索窗口的大小。CamShift 算法可以克服这一缺点,它在对目标定位的同时,还能够自适应调整搜索窗口的大小。CamShift 的算法步骤如下:

(1) 初始化搜索窗口的大小和位置。

(2) 执行 MeanShift 算法。

(3) 按如下方法调整搜索窗口的大小:

首先计算  $B(x, y)$  的二阶矩  $M_{20}$ 、 $M_{02}$ 、 $M_{11}$ :

$$\begin{aligned} M_{20} &= \sum_x \sum_y x^2 B(x, y), \\ M_{02} &= \sum_x \sum_y y^2 B(x, y), \\ M_{11} &= \sum_x \sum_y xy B(x, y) \end{aligned} \quad (4)$$

设  $a = M_{20}/M_{00} - x_c^2$ ,  $b = 2(M_{11}/M_{00} - x_c y_c)$ ,  $c = M_{02}/M_{00} - y_c^2$ 。

然后调整搜索窗口的长和宽:

$$\begin{aligned} l &= \sqrt{(a+c+\sqrt{b^2+(a-c)^2})/2}, \\ w &= \sqrt{(a+c-\sqrt{b^2+(a-c)^2})/2} \end{aligned} \quad (5)$$

(4) 重复步骤(2)和(3),直到搜索窗口收敛。

可以通过将 CamShift 算法扩展到连续的视频图像序列实现对指尖目标的跟踪。它的基本思想是对手指移动视频图像的所有帧的颜色概率分布

的二值化图像作 CamShift 运算,得到搜索目标的位置和大小,并将上一帧的结果作为下一帧 CamShift 算法的搜索窗口的初始值,如此迭代下去,就可以实现对运动指尖目标的跟踪,并得到指尖目标的运动轨迹。由于我们的算法是基于二值图像,大大减少了计算量,加快了算法收敛速度。图 1 显示了指尖跟踪结果。

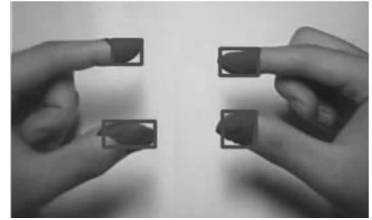


图 1 多个指尖的跟踪效果图

### 2.4 手势的表示

基于机器视觉的手势分成两大类:四指动态手势(双手手势)和两指动态手势(单手手势)。

四指手势由两个手的四个指尖运动轨迹形成。当检测到四个指尖进入拍摄区域是开始手势记录,当四个指尖离开拍摄区域时结束手势记录,并开始手势识别。四指手势主要包括放大、缩小、旋转三种手势。

(1) 放大:左手两个指尖和右手两个指尖相背移动,主要用图像的放大;

(2) 缩小:左手两个指尖和右手两个指尖相向移动,主要用于图像缩小;

(3) 旋转:左手两个指尖和右手两个指尖同时做顺时针或逆时针旋转,主要用于图片旋转。

两指手势由一个手的两个指尖(拇指、食指)运动形成,当检测到有两个指尖渐近且距离小于某个阈值时开始手势记录,当检测到有两个指尖渐远且距离大于某个阈值时结束手势记录。两点手势主要包括 0~9 数字、确认、取消、前进、后退、上翻、下翻共 16 个手势,如表 1 所示。表中的笔画图示的粗圆点表示笔画手势的开始。

表 1 两点手势列表

手势名称	操作语义	笔画图示
数字	输入数字字符	0123L.....
确认	确认/拨号	✓
取消	放弃/取消拨号	✗
前进	前进/向左滚屏	←
后退	后退/向右滚屏	→
下翻	向下翻页/下一标签	↶
上翻	向上翻页/上一标签	↷

对于四指手势,可以通过分析左手两个指尖和右手两个指尖的相对位置变化就可以进行手势判断,而对于两指手势,则需要考虑手势轨迹的表示。

对于一个两指手势,通过指尖跟踪,得到了该手势所经过的点的坐标,构成一个点集。如果点集的点数可能超过 15 个,则查找跨度最小的两个点,用它们的中点取代它们,重复执行,直到手势点集中点的数目为 15 个,由此得到 14 个向量并对它们进行归一化处理。然后将该手势笔画的向量依次连接, $x$ 、 $y$  坐标依次出现,即可构成表示该手势的向量(28 个分量)。比如,对于如图 2 的“确认”和数字“2”手势笔画对应的输入向量分别为:

$\{0.71, 0.71, 0.64, 0.77, 0.67, 0.74, 0.61, 0.79, 0.71, 0.71, 0.88, -0.48, 0.50, -0.86, 0.49, -0.87, 0.51, -0.86, 0.55, -0.83, 0.59, -0.81, 0.54, -0.84, 0.52, -0.85, 0.77, -0.64\}$  和  $\{1.00, 0.00, 1.00, 0.00, 1.00, 0.00, 1.00, 0.00, 0.91, 0.41, -0.22, 0.98, -0.22, 0.97, -0.29, 0.96, -0.27, 0.96, -0.33, 0.94, -0.30, 0.95, -0.21, 0.98, -0.18, 0.98, -0.27, 0.96\}$ 。

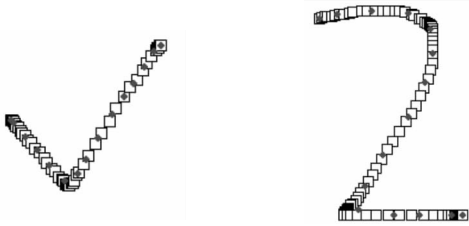


图 2 确认和数字 2 的手势轨迹及其样本点

### 3 MIMO Hermite 神经网络模型

#### 3.1 Hermite 正交多项式

首先给出 Hermite 正交多项式的定义。

定义 1<sup>[14]</sup> Hermite 多项式可直接定义为:

$$\varphi_{h-1}(x) = (-1)^{h-1} e^{x^2} \frac{d^{h-1}(e^{-x^2})}{dx^{h-1}}, 0 \leq x < +\infty \quad (6)$$

它是区间  $[0, +\infty)$  中关于权函数  $\rho(x) = e^{-x^2}$  的  $h-1$  次正交多项式。

Hermite 正交多项式满足如下递推关系式:

$$\varphi_h(x) = (2x\varphi_{h-1}(x) - 2(h-1)\varphi_{h-2}(x)), \quad h = 2, 3, 4, \dots \quad (7)$$

并因此可推算得:

$$\begin{aligned} \varphi_0(x) &= 1 \\ \varphi_1(x) &= 2x \\ \varphi_2(x) &= 4x^2 - 2 \end{aligned}$$

$$\varphi_3(x) = 8x^3 - 12x$$

$$\varphi_4(x) = 16x^4 - 48x^2 + 12$$

.....

依据多项式插值和逼近理论,可用广义多项式函数  $\varphi(x)$  来插值或逼近未知函数  $\Phi(x)$ , 并且对于  $\varphi(x)$  对  $\Phi(x)$  的逼近能力,有如下定义和定理:

定义 2<sup>[15]</sup> (最佳平方逼近) 假设  $\varphi(x)$  和函数系  $\{\varphi_j(x)\}_{j=0}^{h-1}$  中的所有函数  $\varphi_j(x)$  在区间  $[a, b]$  是连续的,且  $\{\varphi_j(x)\}_{j=0}^{h-1}$  是区间  $[a, b]$  上的一个线性无关函数系,  $\rho(x)$  为  $[a, b]$  上的一个权函数。

确定广义多项式  $\varphi(x) = \sum_{j=0}^{h-1} \omega_j \varphi_j(x)$  的系数  $\omega_0, \omega_1, \dots, \omega_{h-1}$ , 使  $\int_a^b [\varphi(x) - \Phi(x)]^2 \rho(x) dx$  最小, 这样得到的函数  $\varphi(x)$  被称为  $\varphi(x)$  在  $[a, b]$  上关于权函数  $\rho(x)$  的最佳平方逼近(函数)。

定理 1<sup>[15]</sup> (最佳逼近之存在与唯一) 设  $\varphi(x)$  在区间  $[a, b]$  是连续的, 则  $\varphi(x)$  的最佳平方逼近函数  $\varphi(x)$  存在而且唯一, 且其系数  $\omega_0, \omega_1, \dots, \omega_{h-1}$  可以通过求解而得。

#### 3.2 MIMO Hermite 神经网络模型和权值确定方法

为适应动态手势识别要求,构造多输入、多输出 Hermite 神经网络模型如图 3 所示,它由输入层、隐含层和输出层组成。其中输入层有  $n$  个节点,  $X = (x_1, x_2, \dots, x_n)^T \in R^n$  为网络的输入向量,输出层有  $m$  个节点,  $Y = (y_1, y_2, \dots, y_m)^T \in R^m$  为网络输出,输出层节点中的  $\Sigma$  表示输出层神经元采用线性激励函数,隐含层有  $h$  个神经元,并采用一组阶次逐渐增高的 Hermite 正交多项式  $\varphi_j(x) (j = 0, 1, 2, \dots, h-1)$  作为其激励函数,  $W \in R^{h \times m}$  为隐含层到输出层的权矩阵,其中  $w_{jk}$  为隐含层第  $j$  个节点与输出层的第  $k$  个节点的连接权值。为了简化模型的网络结构,便于权值学习算法的推导和直接确定,该 Hermite 神经网络中的所有神经元的阈值全部设为 0。

假定有  $P$  个学习样本  $(X^1, X^2, \dots, X^P)$ , 相应的样本输出即教师信号为  $D = (D^1, D^2, \dots, D^P)$ , 为了构建输入输出之间的函数关系,采用基于 Hermite 正交基函数技术进行插值逼近。这些插值函数定义为基函数的线性组合:

$$F_k(X) = \sum_{j=1}^h \omega_{jk} \varphi_j(\|X\|), k = 1, 2, \dots, m \quad (8)$$

其中  $\|X\|$  为  $R^n \rightarrow R$  的线性映射。

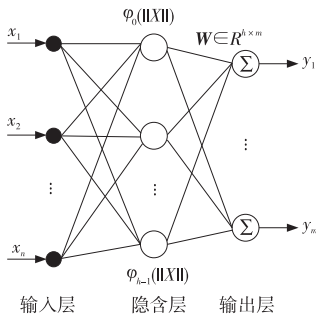


图3 Hermite 正交基前向神经网络模型

借鉴文献[12]的思想,并将其推广到多输入、多输出情况,得到中 Hermite 正交基前向神经网络权值的迭代求取方法和直接确定方法。

**定理 2** (矩阵向量迭代法)图 3 所示的 Hermite 正交基前向神经网络的权值迭代公式的矩阵向量形式可为:

$$\mathbf{W}(k+1) = \mathbf{W}(k) - \eta \boldsymbol{\Phi}^T (\boldsymbol{\Phi} \mathbf{W}(k) - \mathbf{D}) \quad (9)$$

其中:

$$\boldsymbol{\Phi} =$$

$$\begin{bmatrix} \varphi_0(\|X^1\|) & \varphi_1(\|X^1\|) & \cdots & \varphi_{h-1}(\|X^1\|) \\ \varphi_0(\|X^2\|) & \varphi_1(\|X^2\|) & \cdots & \varphi_{h-1}(\|X^2\|) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_0(\|X^P\|) & \varphi_1(\|X^P\|) & \cdots & \varphi_{h-1}(\|X^P\|) \end{bmatrix}$$

$\in R^{P \times h}$  为输入整合矩阵,  $k=0,1,2,\dots$  为迭代次数,学习步长  $\eta$  应足够小以保证迭代收敛。

同传统的 BP 神经网络训练方法一样,定理 2 给出的权值迭代公式需要通过冗长的迭代训练才能获得最优权值。对于这种特殊结构的 Hermite 正交基前向神经网络,可以通过基于伪逆思想的直接确定法一步得到权值,即如下的定理 3。

**定理 3** (权值直接确定)对于图 3 所示 Hermite 正交基前向神经网络的最优权值  $\lim_{k \rightarrow \infty} \mathbf{W}(k)$  可以直接确定:

$$\mathbf{W} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \mathbf{D} \quad (10)$$

其中  $\text{pinv}(\boldsymbol{\Phi}) = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T$  称为  $\boldsymbol{\Phi}$  的伪逆。

**证明** 根据定理 2, 即  $\mathbf{W}(k+1) = \mathbf{W}(k) - \eta \boldsymbol{\Phi}^T (\boldsymbol{\Phi} \mathbf{W}(k) - \mathbf{D})$ , 考虑到学习步长  $\eta > 0$ , 则当网络训练达到稳态后(即  $k$  足够大时)有  $\mathbf{W}(k+1) = \mathbf{W}(k) = \mathbf{W}$ , 即满足  $\boldsymbol{\Phi}^T (\boldsymbol{\Phi} \mathbf{W} - \mathbf{D}) = 0$  的  $\mathbf{W}$  是该神经网络的稳态最优权值  $\lim_{k \rightarrow \infty} \mathbf{W}(k)$ 。因此,可利用伪逆矩阵及公式(12)一步确定最优权值向量。定理 3 因此得证。□

基于伪逆的 Hermite 正交基前向神经网络避免了冗长的迭代训练,同时能够实现输入输出关系的最小平方误差逼近。

### 3.3 网络隐含层结点自适应确定

传统 BP 神经网络隐含层和隐含层节点数目的设定缺乏完善的理论指导,如何根据实际问题来确定具体的网络结构尚无很好的方法,大多依据设计者经验来试凑。若隐含层节点数目设定过少,则无法达到精度要求;设定过多又可能造成隐含层节点冗余,难以确保网络结构最优。

针对上述缺点,在权值直接确定的基础上,设计一种设定隐含层节点自适应算法,能够根据实际问题,快速有效地自动确定网络结构,使得在保证隐含层节点数目最少的同时又能满足目标精度要求。

给定网络目标精度  $\epsilon > 0$ , 满足该精度要求的最小隐含层节点数目自适应确定算法步骤为:

(1) 初始化隐含层节点数目  $h = \sqrt{n+m} + 1$ , 并设定目标精度,如  $\epsilon = 0.005$ ;

(2) 设定神经网络的最大可能隐含层节点数目,如  $\text{MaxHideNode} = \sqrt{n+m} + 20$ ;

(3) 如果  $h > \text{MaxHideNode}$ , 则说明在  $\text{MaxHideNode}$  范围内,无法满足网络目标精度要求,程序退出,否则跳转至步骤(4);

(4) 计算当隐含层节点数目为  $h$  时,神经网络的权值矩阵  $\mathbf{W}$  及最小均方误差(MSE);

(5) 如果  $\text{MSE} \leq \epsilon$ , 则说明找到了满足网络目标精度要求的最小隐含层节点数目  $h$ , 程序退出, 否则  $h = h + 1$ , 跳转至步骤(3)。

上述算法能够快速确定隐含层最小节点数,进而确定神经网络最简化模型结构。

## 4 基于 Hermite 正交基神经网络的动态手势识别

对于四指手势,可以通过分析左后两个指尖和右手两个指尖的相对位置变化就可以进行手势识别,而对于两指手势,则需要基于手势轨迹利用神经网络进行识别。

(1) 神经网络的结构和输入输出设计。

对于图 3 所示的 Hermite 神经网络,输入节点设为 28, 对应前面所说的两指动态手势输入向量的各个分量;隐含层节点数根据设置的网络精度要求 0.005 并采用上述的最小隐含层节点数目自适应确定算法确定为 12;输出层节点数设置为 16, 对应两指动态手势的种类数。由于需要识别 16 种两点手势,所以使用长度为 16 的向量作为网络输出

向量。

### (2) 神经网络的训练。

对于 16 个两指手势,准备  $16 \times 5$  个训练样本,每个手势各取 5 个训练样本。

对于这些训练样本的输入向量,可以通过前述方法获取,而对于这些训练样本的预期输出向量(教师信号),由于两指手势分为 16 种,所以设置 16 个长度为 16 的单位向量对应 16 种手势,其中第  $M$  个元素为 1、其余元素均为 0 的单位向量作为第  $M$  种手势的预期输出向量。例如,1 号手势“0”的预期输出向量为(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),14 号手势“后退”的预期输出向量为(0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0)。

通过将训练样本的输入向量和预期输出向量送入神经网络,通过公式(10),可以求出 Hermite 神经网络的最优权值矩阵  $W$ ,之后该神经网络就可以进行两指动态手势识别。

### (3) 两指动态手势的识别。

通过前述方法获取待识别动态手势的轨迹及向量组合,并以该向量组合作为输入向量输入到已确定最优权值的 Hermite 神经网络,经过运算即可实现手势识别:如果网络输出与第  $M$  种手势的标准输出向量越接近,说明输入手势越接近第  $M$  种手势的标准输入。评判“接近”的简单标准是找出输出向量中最大(最接近 1)的分量,认为输入手势最有可能是该位置为 1 的输出向量对应的手势。例如,输入一个手势向量到神经网络后,输出结果为(0.01,0.00,0.00,0.00,0.01,0.00,0.00,0.00,0.00,0.00,0.99,0.00,0.00,0.00,0.01,0.01),其最大分量是第 11 个分量,我们可以判断该手势为第 11 号手势“确认”手势。

如果输出向量中的最大分量小于某个阈值(如 0.5),则可以判断该手势不是预定义手势,手势识别失败。

## 5 测试结果及分析

为了验证 Hermite 神经网络模型在动态手势学习和识别的可行性和和效果,在训练样本、识别样本相同的条件下,同时采用传统的 BP 神经网络(分别通过带动量梯度下降法 GDX、LM 算法进行训练)和 Hermite 神经网络进行神经网络训练和手势识别,将三者的学习训练情况和手势识别情况进行了对比分析。

### 5.1 手势学习训练

由表 2 可以看出,对于 Hermite 神经网络,采用直接确定法一步确定权值,手势学习训练时间大大缩短,只需要 0.57s,而采用 GDX 和 LM 训练算法的 BP 神经网络,仿真 10 次的平均训练间分别为 15.3s 和 6.5s。同时,采用 BP 神经网络,通过 GDX 和 LM 方法训练后(设定迭代次数上限为 3000 次,网络性能目标为 0.005),网络误差分别为 0.004 992 5 和 0.004 204 1,而采用 Hermite 神经网络,通过直接确定法得出的网络误差为 0.003 462 3。可见,采用 Hermite 正交基前向神经网络相比于 BP 神经网络,不仅在网络训练速度上具有很大的优越性,而且能够达到更好的网络精度,能够更好地逼近动态手势识别这一非线性系统。

表 2 Hermite 神经网络和 BP 神经网络手势学习训练情况对照表

方法	训练迭代次数	训练时间(秒)	网络误差
BP(GDX)	2 213	15.30	0.004 992 5
BP(LM)	197	6.50	0.004 204 1
权值直接确定法	一步计算	0.57	0.003 462 3

### 5.2 手势识别测试

通过上述方法训练完神经网络后,将需要识别的手势样本输入到 Hermite 神经网络进行识别测试。从测试结果可以看出,Hermite 神经网络对于动态手势识别具有较好的鲁棒性和泛化能力,能够容忍手势输入的一定范围的差异,如图 4 所示。

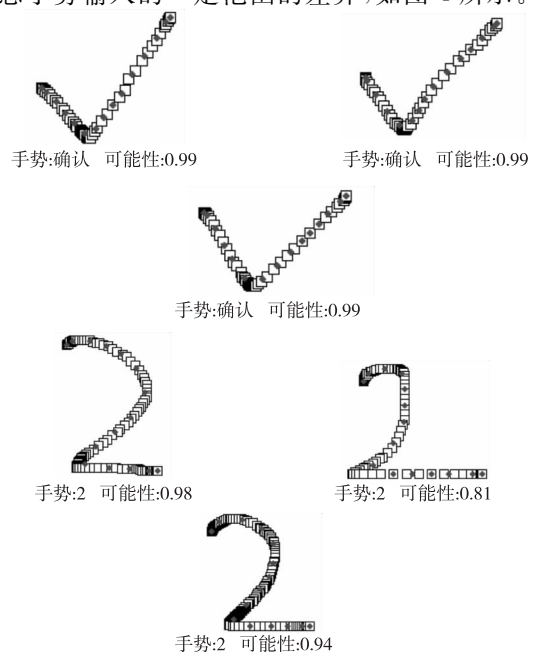


图 4 Hermite 神经网络对于手势输入的鲁棒性和泛化能力

对每一种两指动态手势,分别采集 200 个识别测试样本,然后将识别测试样本输入到已经确定权值的 Hermite 神经网络和 BP 神经网络分别进行识别测试,采用 GDX 和 LM 方法训练的 BP 神经网络的手势识别平均准确率为 89.7% 和 91.9%,而 Hermite 神经网络的手势识别平均准确率达到 94.7%,识别准确率明显得到提高。

## 6 结束语

BP 神经网络在基于机器视觉的手势识别过程中,因网络结构确定缺乏理论指导,训练迭代次数较多,训练速度较慢,不便于用户进行在线的手势学习;同时网络易于过早限于局部收敛,网络训练精度低,影响手势识别准确率。本文将 Hermite 神经网络应用到基于机器视觉的动态手势识别中,通过直接确定网络权值和自适应确定隐含结点数目,大大提高手势的学习训练速度,便于用户进行在线手势学习训练;同时大幅度提高网络的精度和泛化能力,从而提高了网络的动态手势识别能力。

### 参考文献:

- [1] 施彦,韩力群,廉小亲. 神经网络设计与实例分析[M]. 北京:北京邮电大学出版社,2009.
- [2] 刘金琨. 智能控制[M]. 第 2 版. 北京:电子工业出版社,2009.
- [3] 董士海. 人机交互的进展及面临的挑战[J]. 计算机辅助设计与图形学学报,2004,16(1):1-11.
- [4] Moya J M, de Espinosa A M, Araujo A, et al. Low-Cost Gesture-Based Interaction for Intelligent Environments[C]// Proc of IWANN 2009, 2009:752-755.
- [5] Alexander T C, Ahmed H S, Anagnostopoulos G C. An Open Source Framework for Real-Time, Incremental, Static and Dynamic Hand Gesture Learning and Recognition[C]// Proc of HCI'09, 2009:123-130.
- [6] Kaustubh, Sumantra. Hand Gesture Modelling and Recognition Involving Changing Shapes and Trajectories, Using a Predictive EigenTracker[J]. Pattern Recognition Letters, 2007, 28:329-334.
- [7] Francke H, Ruiz-del-Solar J, Verschae R. Real-Time Hand Gesture Detection and Recognition Using Boosted Classifiers and Active Learning[C]// Proc of PSIVT'07, 2007:533-547.
- [8] 施教芳,陈珏,赵鹤鸣. 基于节点对 BP 网络的数据手套手势识别[J]. 系统仿真技术,2008,4(3):154-157.

- [9] 肖志勇,秦华标. 基于视线跟踪和手势识别的人机交互[J]. 计算机工程,2009,35(15):198-200.
- [10] 吴小俊,王士同,杨静宇,等. 基于正交多项式函数的神经网络及其性质研究[J]. 计算机工程与应用,2002,38(9):25-26.
- [11] 曹飞龙,张永全,潘星. 构造前向神经网络逼近多项式函数[J]. 模式识别与人工智能,2007,20(3):331-335.
- [12] 张雨浓,肖秀春,陈扬文,等. Hermite 前向神经网络隐节点数目自动确定[J]. 浙江大学学报(工学版),2010,44(2):271-275.
- [13] 徐琨,贺昱曜,王卫亚. 基于 CamShift 的自适应颜色空间目标跟踪算法[J],计算机应用,2009,29(3):757-760.
- [14] 林成森. 数值分析[M]. 北京:科学出版社,2007.
- [15] 莫国端,刘开第. 函数逼近论方法[M]. 北京:科学出版社,2003.



李文生(1966-),男,湖南郴州人,硕士,副教授,研究方向为嵌入式软件开发、多媒体处理与通信。E-mail: lws7166@126.com

LI Wen-sheng, born in 1966, MS, associate professor, his research interests include embedded software, multimedia processing and communications.



解梅(1955-),女,四川成都人,博士,教授,博士生导师,研究方向为数字图像处理 and 生物特征识别。E-mail: mxie@uestc.edu.cn

XIE Mei, born in 1955, PhD, professor, PhD supervisor, her research interests include digital image processing, and biometrics.



邓春健(1980-),男,广东韶关人,博士,副教授,研究方向为信息显示技术和通信技术。E-mail: dcj5880870@126.com

DENG Chun-jian, born in 1980, PhD, associate professor, his research interests include information display technology, and communication technology.



姚琼(1977-),女,江西南昌人,硕士,讲师,研究方向数字图像处理和模式识别。E-mail:2542807@qq.com

YAO Qiong, born in 1977, MS, lecturer, her research interests include digital image processing, and pattern recognition.