

基于 WT 和 LVQ 网络的多姿态人脸识别

陈 蕾, 黄贤武, 孙 兵

(苏州大学电子信息学院, 苏州 215021)

摘 要: 提出了基于小波变换和学习矢量量化网络相结合的新方法进行人脸识别。小波变换具有良好的多尺度特征表达能力, 能将图像的大部分能量集中到最低分辨率子图像, 可以很好地对图像降维和表征人脸图像的特征。LVQ 算法是在有教师状态下对竞争层进行训练的一种学习算法。LVQ 网络结构简单, 但却表现出比 BP 网络更强的有效性和鲁棒性。实验表明该方法对表情和姿态变化的人脸具有良好的分类性能和识别效率。

关键词: 小波变换; 学习矢量量化; 神经网络; 分类; 多姿态人脸识别

Pose-varied Face Recognition Based on WT and LVQ Network

CHEN Lei, HUANG Xianwu, SUN Bing

(School of Electronics and Information Engineering, Soochow University, Suzhou 215021)

【Abstract】 A method of pose-varied face recognition based on wavelet transform and LVQ network is proposed. Wavelet transform has a good ability to express the multi-resolution characteristics. The most energy of an image is mainly concentrated on its lowest resolution subimage after wavelet transform. So it can reduce the dimension and extract the main features of an image. Learning vector quantization (LVQ) is an effective learning algorithm that trains the competitive layer under supervision. It has a simple network structure, but it is very effective and robust in face recognition. The experiment result shows that the method has very good classification capability and high recognition rate for pose-varied faces.

【Key words】 Wavelet transform; Learning vector quantization(LVQ); Neural network; Classification; Pose-varied face recognition

人脸识别作为图像处理和模式识别领域的研究热点, 有非常广泛的应用前景, 如身份识别、智能门卫等。在现有的许多识别方法中, 识别率是最值得关注的问题。由于人脸十分相似, 寻找有效的特征是识别率的关键。传统的 PCA 方法可以有效地降维, 但它是一种线性算法, 不能抽取出数据中的非线性特征, 因此会影响识别率, 并且计算量大。

被誉为数学显微镜的小波变换具有多分辨率分析的特点, 在时频两域都具有表征信号局部特征的能力。图像经过小波变换, 其低频部分保留了绝大部分信息和能量。利用这个特点可以有效地对图像特征进行压缩和提取。

学习矢量量化网络是一种混和网络, 由输入层、竞争层和线性输出层组成。通过有监督及无监督的学习进行分类。比起普遍使用的 BP 网络, 有其独特的优势。BP 网络的缺点在于采用基于梯度下降的非线性优化策略, 有可能陷入局部最小问题, 不能保证求出全局最小值。其它一些优化策略如遗传算法、模拟退火算法等, 虽然可以求得全局最小, 但是计算量很大, 有时候会出现效率问题。LVQ 网络的优点是不需要将输入向量进行归一化、正交化, 只需要直接计算输入向量与竞争层之间的距离, 即可实现模式识别, 识别效率高, 而且对姿态变化的人脸图像具有一定的鲁棒性。

1 图像的小波变换降维

原始图像信息量较大, 直接用来识别会影响识别速度, 因此需要压缩。小波变换通过多分辨率分析过程将一幅图像分成近似和细节 2 部分, 细节对应的是小尺度的瞬变, 它在本尺度内很稳定。因此将细节存储起来, 对近似部分在下一个尺度上进行分解, 重复该过程即可得到多尺度图像, 如图 1 所示。近似与细节在正交镜像滤波器算法中分别对应于高

通和低通滤波, 这种变换通过尺度去掉相关性, 在图像压缩中非常有效。

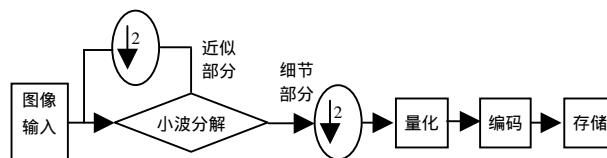


图 1 基于小波变换的图像压缩过程

对图像进行一次二维小波分解后, 可得到 4 个区域: 低频区域是近似分量, 高频区域分别是水平分量、垂直分量和对角分量。每次变换后的近似分量可以再次进行小波变换。图 2 为经过两层小波变换压缩后的结果。

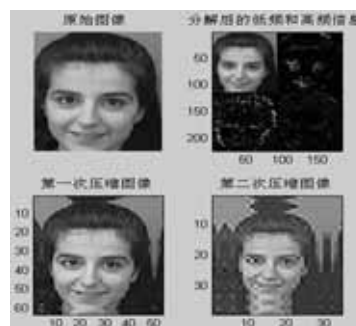


图 2 基于小波变换的图像压缩

表 1 中是压缩前后的结果, 可以看出, 随着分解层数的增加, 压缩比是递减的。

作者简介: 陈 蕾(1967 -), 女, 博士生, 主研方向: 图像处理与模式识别; 黄贤武, 教授、博导; 孙 兵, 副教授

收稿日期: 2006-01-25 **E-mail:** chenlei@suda.edu.cn

表 1 图像压缩结果

Name	Size	Bytes	Class
压缩前图像 X	112x92	82 432	double array
第 1 次压缩后图像 ca1	63x53	26 712	double array
第 2 次压缩后图像 ca2	39x34	10 608	double array

2 LVQ 网络

2.1 网络模型

LVQ 网络模型如图 3 所示。一个 LVQ 网络由 3 层神经元组成，即输入层、竞争层、输出层。该网络在输入层与竞争层之间为完全连接，而在竞争层与输出层之间为部分连接，每个输出神经元与竞争神经元的不同组相连接。竞争层和输出神经元之间的连接权值固定为 1。输入层和竞争神经元间的连接权值建立为参考矢量的分量(对每个竞争神经元指定一个参考矢量)。在网络训练过程中，这些权值被修改。竞争神经元和输出神经元都具有二进制输出值。当某个输入模式被送至网络时，参考矢量最接近输入模式的竞争神经元因获得激发而赢得竞争，因而允许它产生一个“1”。其他竞争神经元都被迫产生“0”。与包括获胜神经元的竞争神经元组相连接的输出神经元也发出“1”，而其他输出神经元均发出“0”。产生“1”的输出神经元给出输入模式的类，每个输出神经元被表示为不同的类。

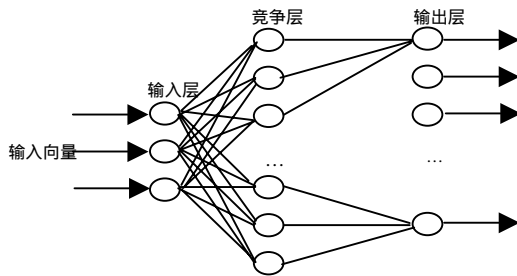


图 3 LVQ 网络模型

2.2 LVQ 网络学习规则

利用 LVQ 网络实现模式识别，不需要将输入向量归一化、正交化，只需要直接计算输入向量与竞争层之间的距离，从而实现识别。

假定网络输入层的输入向量为 $X=(x_1, x_2, \dots, x_M)$ ，其中， M 为输入神经元的数目；输入层和竞争层之间的连接权值矩阵为 $W^1=(w_1^1, w_2^1, \dots, w_P^1)$ ， $w_i^1=(w_{i1}^1, w_{i2}^1, \dots, w_{iM}^1)$ 。 w_{ij}^1 中， $i=1, 2, \dots, P$ ， $j=1, 2, \dots, M$ ，表示输入层第 i 个神经元和第 j 个神经元间的连接权值， P 为竞争神经元的数目，竞争层的输出向量为 $V=(v_1, v_2, \dots, v_P)$ ，竞争层与输出层神经元间的连接权值矩阵为 $W^2=(w_1^2, w_2^2, \dots, w_N^2)$ ，其中 $w_k^2=(w_{k1}^2, w_{k2}^2, \dots, w_{kp}^2)$ 。 w_{kr}^2 中， $k=1, 2, \dots, N$ ， $r=1, 2, \dots, p$ ，表示竞争层第 k 个神经元与输出层第 r 个神经元之间的连接权值， N 为输出层神经元的数目。竞争层的每个神经元通过学习原型向量，并对输入空间进行分类。将竞争层学习得到的类称为子类，将输出层学习得到的类称为目标类。

LVQ 网络的学习结合了竞争学习和有监督学习的规则。正如所有有监督学习算法一样，它需要一组正确网络行为的例子来训练网络：

$$\{x_1, t_1\}, \{x_2, t_2\}, \dots, \{x_Q, t_Q\}$$

其中，每个目标输出向量 t_j ($j=1, 2, \dots, Q$) 有且只有一个分量为 1，其它分量全都为 0。通常，把竞争层的每个神经元指定给一个输出神经元，这样就可以定义矩阵 W^2 。 W^2 的列表示子类，行表示类。 W^2 的每一列仅有一个 1 出现的行，表明这个子类

属于该行表示的类，即

$$W_{kr}^2 = \begin{cases} 1 & \text{有且仅有 } r \in k \\ 0 & \text{如果有 } r \in k \end{cases}$$

W^2 一旦定义好就不再改变，神经网络的学习是通过改进了的 Kohonen 规则使 W^1 改变进行的，LVQ 学习算法为：

第 1 步 设置变量和参数： $X(n)=[x_1(n), x_2(n), \dots, x_N(n)]^T$ 为输入向量，或称训练样本。

$W_{i1}(n)=[w_{i1}(n), w_{i2}(n), \dots, w_{iN}(n)]^T$ 为权值向量， $i=1, 2, \dots, M$ 。选择学习速率的函数 $\eta(n)$ 。 n 为迭代次数， N 为迭代总次数。

第 2 步 初始化权值向量 $W_i(0)$ 及学习速率 $\eta(0)$ 。

第 3 步 从训练集中选取输入向量 X 。

第 4 步 通过欧氏距离最小的标准：

$$\|X - W_c\| = \min_i \|X - W_i\|, i=1, 2, \dots, M \quad (1)$$

寻找获胜神经元 c ，从而实现了神经元的竞争过程。

第 5 步 判断分类是否正确，根据如下规则调整获胜神经元的权值向量：

用 L_{W_c} 代表与获胜神经元权值向量联系的类，用 L_{X_i} 代表与输入向量相联系的类。

如果 $L_{X_i} = L_{W_c}$ ，则

$$W_c(n+1) = W_c(n) + \eta(n)[X - W_c(n)] \quad (2)$$

否则，当 $L_{X_i} \neq L_{W_c}$ ，有

$$W_c(n+1) = W_c(n) - \eta(n)[X - W_c(n)] \quad (3)$$

对于其它神经元，保持权值不变。

第 6 步 调整学习速率 $\eta(n)$

$$\eta(n) = \eta(0)(1 - \frac{n}{N}) \quad (4)$$

第 7 步 判断迭代次数是否超过 N ，如果 $n > N$ 就转到第 3 步，否则结束迭代过程。

2.3 LVQ 的改进算法 LVQ2

学习矢量化作为有监督学习指定了输出神经元代表的类，而在用随机值对网络的权值进行初始化时并没有考虑输出层中指定类别的分布，这个矛盾可能导致权值向量不能达到目标，从而不能准确分类。改进算法 LVQ2 的主要思想是：通过引入一个次获胜神经元来增加获胜权值训练的神经元个数，从而缓解学习矢量化量化的内在矛盾，使网络实现正确的分类。

LVQ2 的改进主要在网络的训练方面，通过训练获胜神经元和次获胜神经元来扩展学习矢量化。所谓次获胜神经元满足以下条件：该神经元的权值向量和输入神经元的权值向量的欧氏距离，仅大于获胜神经元的权值向量和输入神经元的欧氏距离。算法的主要思想是：当获胜神经元 C 不代表正确分类而次获胜神经元 R 却代表正确分类时，对二者的权值均进行训练。对获胜神经元进行惩罚，而对次获胜神经元 R 进行奖励。LVQ2 只有在输入向量到获胜神经元 C 的权值向量的距离 d_C 和到次获胜神经元 R 的权值向量的距离 d_R 没有很大的差别时，才和原有的 LVQ 有区别，否则和 LVQ 相同。在此将 d_C 和 d_R 之间没有多大差别的条件用一个“窗”来表示，定义为

$$\frac{d_C}{d_R} = 1 - \varepsilon \quad (5)$$

其中， ε 为给定阈值。但满足该条件时，用以下方法调整

权值：

$$\begin{aligned} W_R(n+1) &= W_R(n) + \eta[X - W_R(n)] \\ W_C(n+1) &= W_C(n) - \eta[X - W_C(n)] \end{aligned} \quad (6)$$

其中， W_R 是次获胜者 R 的权值向量，它和训量向量同属一类，而 W_C 为获胜者 C 的权值向量。

3 基于 LVQ 网络的多姿态人脸识别

(1) 输入向量和目标向量的设计

从 ORL 人脸数据库中随机选取 10 人共 100 幅不同姿态的人脸图像，50 幅用于训练，50 幅用于测试。每幅图像的原始尺寸为 112×92 ，经过 2 次小波变换降维成 39×34 ，再利用最临近插值原理统一成 25×20 ，经过行堆叠人脸图像变成 1×500 的向量。由于参与训练的有 50 幅图像，因此输入向量为 500×50 。50 幅人脸图像分为 10 类，所以目标向量为 10×50 的向量。其中每列中只有 1 个“1”，其余均为“0”。

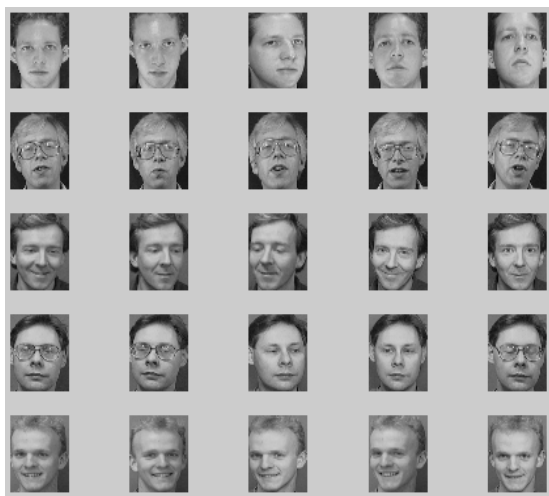


图 4 用于训练的部分图像

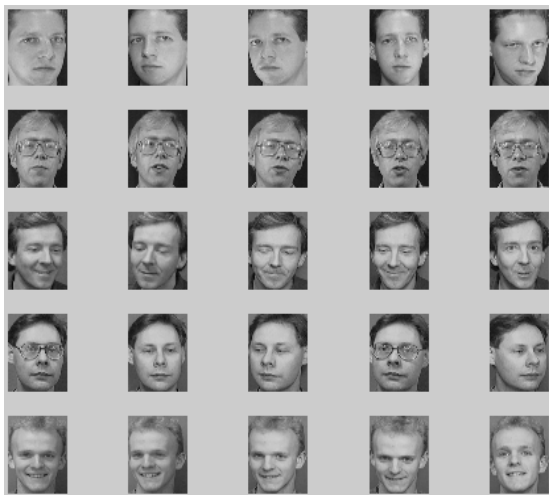


图 5 用于测试的部分图像

(2) 网络创建、训练和测试

首先创建一个两层网络。网络有 500 个输入节点，10 个输出神经元。中间层节点数目的选择是一个复杂的问题。数目过少，网络将不能建立复杂的判断界，网络不能训练出来，容错性差；数目过多，会使学习时间过长。实际上，精确预测该层的神经元数目至今仍然存在理论上还没有解决的问题。对于 LVQ 网络，竞争获胜的神经元表示的是一个子类而非一个类，一个类可能由几个不同的神经元(子类)组成。典型情况下，相同数量的隐层神经元连接到每个输出神经元，

由于输出神经元有 10 个，因此可以初步设定中间层的神经元个数为 20，即每个类是两个子类的联合。

对连接权值初始化。在初始状态下，所有神经元都拥有相同的权值，即这些向量的中间值。学习速率反映了学习过程中连接权调整量的大小，初值设为 0.1。

最后是期望误差的选取。在设计网络的训练过程中，期望误差值应当通过对比训练后确定一个合适的值，较小的期望误差值要靠增加隐含层的节点以及训练时间来获得。本文期望误差取值为 0.1。

学习函数用 $lvq2$ ，最大训练步数设为 500，开始对网络进行训练。训练过程曲线如图 6 所示。

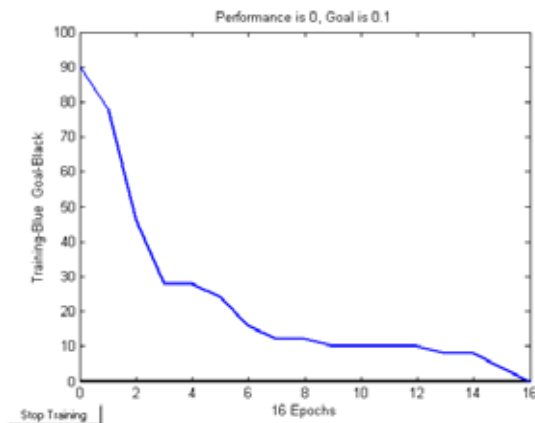


图 6 误差容限为 0.1 时的迭代次数 - 训练误差曲线

从图 6 中可以看到，收敛速度很快，经过 16 步就可达到误差要求。训练好的网络用 50 幅图像测试的结果是：当竞争层节点数为 10 时，不能有效识别；当竞争层节点数为 20 时，识别率为 96%；当竞争层节点数为 30 时，识别率为 98%。竞争层节点数再增加，对识别结果影响不大。训练样本的选取对测试结果影响非常大，若有目的地选取训练样本，识别率还可提高。

4 结论

本文提出了基于 LVQ 网络的多姿态人脸识别的新方法。它首先对图像进行小波变换，经过小波变换的图像被送到 LVQ 网络进行训练与测试。实验证明 LVQ 网络能够发挥竞争学习和有监督学习二者的优点，有效地提高了识别率，而且实现简单、速度快。

参考文献

- 1 Hagan M T, Demuth H B, Beale M H. 神经网络设计[M]. 北京: 机械工业出版社, 2002.
- 2 褚蕾蕾. 计算智能的数学基础[M]. 北京: 科学出版社, 2002.
- 3 边肇祺, 张学工. 模式识别(第 2 版)[M]. 北京: 清华大学出版社, 2000.
- 4 黄贤武, 仲兴荣, 王加俊. 基于 WT/PCA 的自适应神经网络人脸识别方法[J]. 计算机应用, 2003, 23(6).
- 5 Aitkenhead M J, McDonald A J S. A Neural Network Face Recognition System[J]. Engineering Applications of Artificial Intelligence, 2003, 16(3): 167-176.
- 6 Javad H, Majid A. N-feature Neural Network Human Face Recognition[J]. Image and Vision Computing, 2004, 22(12): 1071.
- 7 Khuwaja G A. An Adaptive Combined Classifier System for Invariant Face Recognition[J]. Digital Signal Processing, 2002, 12(1): 21-46.