

# 基于深度学习与融入梯度信息的人脸姿态分类检测

苏铁明<sup>1</sup> 程福运<sup>1</sup> 韩兆翠<sup>2</sup> 欧宗瑛<sup>1</sup>

(1. 大连理工大学机械工程学院, 大连, 116024; 2. 济南大学机械工程学院, 济南, 250000)

**摘要:** 针对人脸姿态分类问题, 本文提出了一种基于深度学习与融入梯度信息的人脸姿态分类学习方法。首先提取人脸姿态图像灰度与灰度差组合特征, 然后通过三层受限玻尔兹曼机(Restricted Boltzmann machines, RBM)对大量样本的特征进行融合训练学习, 提取反映人脸姿态内涵的深度学习特征。最后通过 Softmax 分类器建立深度学习特征与人脸姿态标签的对应关系。在对 CAS-PEAL-R1 人脸数据库进行学习和分类检测中, 获得普遍高于 95% 的分类精度。

**关键词:** 人脸姿态分类; 深度学习; 受限玻尔兹曼机

**中图分类号:** TP391      **文献标志码:** A

## Pose Classification of Human Face Based on Deep Learning and Gradient Information Fusion

Su Tieming<sup>1</sup>, Cheng Fuyun<sup>1</sup>, Han Zhaocui<sup>2</sup>, Ou Zongying<sup>1</sup>

(1. School of Mechanical Engineering, Dalian University of Technology, Dalian, 116024, China; 2. School of Mechanical Engineering, University of Jinan, Jinan, 250000, China)

**Abstract:** Aiming at upgrading the performance of face pose classification, we proposed an algorithm of face pose classification based on deep learning and gradient information fusion. First, the pixel gray intensity features and the features of gray intensity difference nearby each pixel from a face image are extracted. Then, these features of face images are processed with deep learning technique through a dedicated three-layer restricted Boltzmann machines network, which has been trained by a large number of samples. Finally, a corresponding relation between fusion deep learning features and the labels of face pose classifications is built through a Softmax classifier. The experiment results show that the proposed algorithm achieves a state of the art classification accuracy, generally higher than 95%, when learning and testing on CAS-PEAL-R1 face database.

**Key words:** face pose classification; deep learning; restricted Boltzmann machines

## 引 言

人脸姿态检测, 即人脸在三维空间的方位角度测定, 是计算机视觉和模式识别领域的一个重要课题。人脸姿态朝向反映了一个人关注的焦点, 也在一定程度上反映了人的意图和情感, 判别人脸姿态

朝向在智能人机交互、计算机游戏、虚拟现实、驾驶员疲劳状态实时监测等领域都有直接的应用。人脸姿态检测是许多智能处理,如人脸身份识别<sup>[1]</sup>、人脸图像的计算机智能分析和理解等的基础。对姿态检测的处理计算,既有精度要求,又有实时检测速度要求,是一个既有重要实用价值又有重要理论意义的热门研究课题<sup>[2]</sup>。许多学者和研究人员在此领域工作多年,虽然已取得了一定进展,但仍是一项有难度的研究课题。人脸姿态检测方法可简单归纳为3大类:(1)基于人脸某些特征点(主要是两眼中心、鼻尖和嘴巴中心等)相对位置模板的判别法<sup>[3]</sup>。这类方法原理简单直观,但准确判别特征点的位置本身就是一个研究课题。特征点位置偏差对匹配结果影响很大,而且姿态方位与模板对应关系也很难做到精确。同时该方法对人脸的仰俯姿态变化尤其不敏感。这类方法对人脸姿态的判别效果并不理想,近年来发表的文章中多已不采用此类方法;(2)基于仿射变换重构三维模型的判别法<sup>[4]</sup>。基于三维模型分析姿态,理论上似乎更合理,但由二维图像重构三维模型本身就是一个难题,只有单张或很少几张人脸照片重构三维模型,又是一个欠约束问题,存在许多不确定性,受这些问题的限制,这类方法近期难以达到实用;(3)基于人脸图像数字特征,通过统计学习分类的判别法。这类方法是当前的主流,也已取得了较好的判别效果。早期此类方法有主成分分析<sup>[5]</sup>、支持向量机(Support vector machine, SVM)<sup>[6]</sup>、弹性图匹配<sup>[7]</sup>及独立成分分析<sup>[8]</sup>等方法。文献[9]提出一种基于局部 Gabor 二值模式(Local Gabor binary pattern, LGBP)的人脸姿态估计算法。该方法先对原图像提取 Gabor 特征图像,然后在 Gabor 图像上提取局部二值模式特征,最后用 SVM 方法进行分类。虽然该方法在识别率上取得了较好效果,但计算量太大,难以满足实时性的要求。文献[10]提出了一种基于距离转化和归一化互相关的人脸姿态估测方法,达到了当时最高的正确识别率。文献[11]采用优选尺度的曲波特征组合,采用非线性的局部线性嵌入(Locally linear embedding, LLE)的方法提取特征,该方法对光照、表情和部分遮挡等不利条件具有较好的鲁棒性,取得了比文献[10]更优的姿态判别效果。

基于数字特征识别姿态精度提高的关键是数字特征的选取。上述已有方法在构建数字特征中存在以下不足:首先所采用的数字特征完全由开发者人工选定。这些特征虽然在其他对象识别应用中证明有效,但对人脸姿态识别应用未必最合适,可能需要作进一步优化适配;其次所采用的特征都是只经过单次处理生成的浅层特征。研究表明,对于复杂的对象识别需要构建包含层次结构的复杂特征才能有效地区分分类。近年兴起的深度学习<sup>[12,13]</sup>理论及技术,通过对大数据的分析和学习,为依靠机器智能自动提取特征,提供了有力的原理和方法,并已在多个领域取得了成效<sup>[14-17]</sup>。常规的深度学习图像分析识别方法只取图像像素灰度值作为输入,通过多层学习自动生成深度学习特征。鉴于图像的梯度信息对于描述人脸各部轮廓的变异具有很强的区分性。本文取灰度值和梯度信息共同输入到专用的深度学习系统,通过融合学习构建深度学习特征,用于姿态检测分类,取得了较高的正确检测精度。

## 1 受限玻尔兹曼机与 Softmax 分类器

深度学习网络系统由受限玻尔兹曼机(Restricted Boltzmann machines, RBM)和 Softmax 分类器两大基本模块构建。

### 1.1 受限玻尔兹曼机

RBM 是由玻尔兹曼机(Boltzmann machines, BM)改进而来的一种产生式概率人工神经网络模型。它由可见层  $v$  和隐藏层  $h$  组成,设可见层单元数为  $m$ ,隐藏层单元数为  $n$ 。同层内各单元无连接,可见层单元  $v_i$  与隐藏层单元  $h_j$  之间通过网络参数:连接系数  $W=[W_{ij}]m \times n$ 、可见层偏置  $a=[a_i](i=1 \sim m)$  和隐藏层偏置  $b=[b_j](j=1 \sim n)$  进行连接,RBM 网络结构如图 1 所示。

对于可见层单元与隐藏层单元均为二值型的标准型RBM,当既定的网络参数以及可见层的输入数据与隐藏层的特征数据均已知时,该RBM系统的能量为

$$e(\mathbf{v}, \mathbf{h}) = - \sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j v_i w_{ij} h_j \quad (1)$$

式中:  $v_i$  为可见层输入数据;  $h_j$  为隐藏层特征数据。根据式(1)定义的能量函数,可以得出可见层单元与隐藏层单元的联合概率为

$$p(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v}, \mathbf{h})} \quad (2)$$

式中:  $Z = \sum_i \sum_j e^{-E(v_i, h_j)}$ 。根据式(2)的联合概率函数,可以推导出可见层的边缘概率为

$$P(\mathbf{v}) = \frac{1}{Z} \sum_h e^{-E(\mathbf{v}, \mathbf{h})} \quad (3)$$

当网络参数和可见层的输入数据已知时,则隐藏层的第  $j$  个单元被激活(即取值为1)的概率  $p(h_j = 1 | \mathbf{v})$  为

$$p(h_j = 1 | \mathbf{v}) = \text{sigmoid}(b_j + \sum_i \omega_{ji} v_i) \quad (4)$$

式中:  $\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$ 。当网络参数和隐藏层特征数据已知时,重构可见层单元,则可见层的第  $i$  个单元被激活(即取值为1)的概率  $p(v_i = 1 | \mathbf{h})$  为

$$p(v_i = 1 | \mathbf{h}) = \text{sigmoid}(a_i + \sum_j \omega_{ji} h_j) \quad (5)$$

对于给定的训练样本集  $\mathbf{V}$ , 每个实例为  $\mathbf{v}$ , RBM 的训练就是以最大化每个实例的概率  $P(\mathbf{v})$  的乘积为目标函数,进行优化求解

$$\arg \max_{\mathbf{w}} \prod_{\mathbf{v} \in \mathbf{V}} p(\mathbf{v}) \quad (6)$$

RBM 的训练过程即网络参数的优化过程。优化目标是使得在优化网络参数下隐藏层所重构的可见层数据概率分布尽可能地与训练数据相符合。隐藏层所提取的特征数据就可以看作样本另一种表达形式的特征,即单层RBM模型最优的输出。2002年, Hinton<sup>[18]</sup>提出了一种使用对比散度的RBM快速学习算法,有效地实现了上述目标函数求解。2006年, Hinton<sup>[12]</sup>在《Science》上发表的论文发表了采用多层RBM取得的成果,受到了机器学习领域研究者广泛的关注。采用RBM网络构建数字特征表达,是当前机器学习领域数字建模中得到普遍认可的技术<sup>[13]</sup>。

深层分类网络由多层RBM与分类器叠加组成。底层RBM以数据样本特征为输入,经过该层RBM学习后输出特征则作为次层的输入,继续进入次层RBM学习。经过多层学习,其顶层的输出则作为深度学习的最后输出特征进入分类器分类处理。

### 1.2 Softmax 分类器

Softmax 是一种有成效、有监督的学习多类别分类器。假设已知训练样本个数为  $m$ , 样本特征维数为  $n$ , 样本分类数目为  $k$ , 则对于任意样本属于第  $j$  类 ( $1 \leq j \leq k$ ) 的概率为

$$p(y_i = j) = \frac{e^{\theta_j \cdot \mathbf{x}_i}}{\sum_{d=1}^{d=k} e^{\theta_d \cdot \mathbf{x}_i}} \quad (7)$$

式中:  $\mathbf{x}_i$  为样本所对应的特征向量;  $\theta_j$  为第  $j$  类所对应的  $n$  维参数向量。  $p(y_i = j)$  为最大值对应的类别  $j$  就是检测的分类结果。式(8)为 Softmax 的代价函数,通过迭代的优化算法(梯度下降法)训练参数,使代价函数值最小,即有

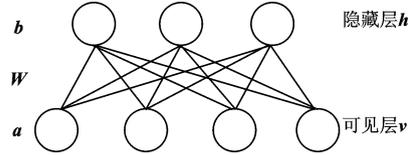


图1 RBM模型  
Fig.1 Model of RBM

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k L\{y(i)=j\} \log \frac{e^{\theta_j^T x_i}}{\sum_{d=1}^k e^{\theta_d^T x_i}} \right] + \frac{\lambda}{2} \sum_{i=1}^m \sum_{j=1}^n \theta_{ij}^2 \quad (8)$$

式中:  $L\{\}$  为示性函数, 其取值规则为:  $L\{\text{表达式正确}\} = 1, L\{\text{表达式错误}\} = 0$ ; 代价函数的第 2 项为防止过拟合项;  $\lambda$  为防止过拟合的系数。

## 2 输入数据预处理

待测试图像经过 Viola-Jones 法<sup>[19]</sup> 对人脸进行检测和初步定位后, 缩放并剪裁为  $28 \times 28$  像素作为姿态检测对象。RGB 彩色图像包含了 3 种通道的不同信息, 但 RGB 各通道值受光照影响较大, 每个通道的值不稳定。将 RGB 图像转化为灰度图像, 可保留图像的亮度信息和轮廓信息, 也可以减少图像数据的维度。本文通过  $Y = 0.299 \times R + 0.587 \times G + 0.114 \times B$  将彩色图像转化成灰度图像 ( $Y$  为灰度图像值,  $R, G, B$  分别代表红、绿、蓝通道值), 然后提取以下特征数据:

(1) 像素灰度特征  $I$ 。将图像的灰度值除以 255, 归一化为  $0 \sim 1$  之间的实数作为网络的输入灰度特征; (2) 图像近邻像素的灰度差特征  $d_1, d_2$ 。灰度差特征可有效地捕获图像中轮廓和纹理信息, 而不同的人脸姿态图像中各部轮廓分布有明显的不同, 因此灰度差特征的分布与人脸姿态的变化具有很强的相关性。

在一个像素点的灰度梯度可以用两个正交方向的梯度值综合表示。本文取为两斜正交方向, 而梯度可以用固定距离像素的灰度差表征。

如图 2 所示, 像素  $A$  两正交斜向的邻域像素为  $a, b, c, d$ 。此两方向邻域的灰度差  $\Delta_1$  和  $\Delta_2$  定义为

$$\begin{aligned} \Delta_1 &= I_a - I_c \\ \Delta_2 &= I_d - I_b \end{aligned} \quad (9)$$

式中:  $I_a, I_b, I_c, I_d$  分别为  $a, b, c, d$  像素的灰度值, 像素值变动范围为  $0 \sim 255$ 。  $\Delta_1$  和  $\Delta_2$  的变动范围为  $-255 \sim 255$ 。使用 Sigmoid 函数, 将  $\Delta_1$  和  $\Delta_2$  归一化为  $0 \sim 1$  的实数  $d_1$  和  $d_2$ , 有

$$d_1 = \frac{1}{1 + e^{-\frac{\Delta_1}{40}}} \quad (10)$$

$$d_2 = \frac{1}{1 + e^{-\frac{\Delta_2}{40}}} \quad (11)$$

式中: 常数 40 为根据测试实践效果选取的经验数值。

一幅  $28 \times 28$  人脸图像共有 784 个像素, 可提取 784 维灰度特征和 1 568 维灰度差特征组成 2 352 维特征向量。从信息源分析, 对于一幅图像其全部像素灰度特征已包含了全部信息。灰度差特征只是由灰度特征派生生成的特征, 并没有增添新的信息内容, 但是它可能揭示了全部信息中有较大意义内涵成分。在本文研发实践中, 取灰度和灰度差组合特征作为输入数据可以取得比只取灰度数据作输入数据更高的检测分类精度。人工优选能力与深度机器学习自动优化能力的结合, 可以比单纯机器学习取得更好的性能。

## 3 人脸姿态分类检测深度网络系统

人脸姿态分类检测网络系统由多层 RBM 网络与 Softmax 分类器叠加而成。

### 3.1 人脸姿态转角和分类

在笛卡尔坐标系中, 人脸在三维空间中的姿态可以分解为 3 个转动, 即绕  $X$  轴的仰俯 (Pitch) 转动、绕  $Y$  轴的旋转 (Roll) 转动及绕  $Z$  轴的偏航 (Yaw) 转动, 如图 3 所示。人脸姿态可以用转动角度定

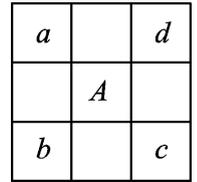


图 2 图像像素邻域关系图  
Fig. 2 Diagram of pixel neighborhood of images

量表示,其中绕Y轴的Roll角度可以通过两眼中心的连线与水平或竖直方向的夹角确定,比较容易判定。本文只研讨对Pitch角度和Yaw角度进行检测,在检测中只按预先定义的被测角度分类范围判定其所属分类。

根据收集到的人脸姿态样本,本文分别对3种类别姿态分类进行深度学习和检测。第1种类别是只按Yaw角度分类,即不区分Pitch角度(仰俯姿态)的变化,在检测中把人脸姿态分为7类,如图4所示;第2种类别是只按Pitch角度分类,即不区分Yaw角度(偏航姿态)的变化,把人脸姿态分为抬头、平视和低头;第3种类别是同时区分Yaw角度和Pitch角度的变化,检测中将姿态样本分为21类,如图5所示。

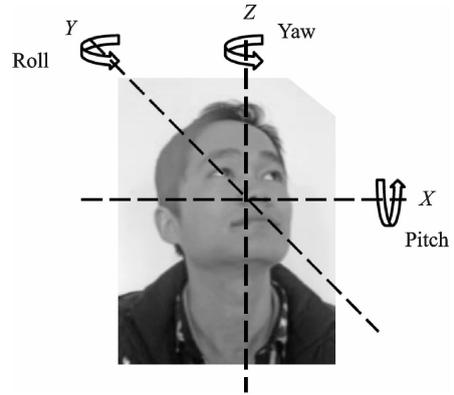


图3 人脸在笛卡尔坐标系下方位变化示意图  
Fig. 3 Sketch map of face orientations change under the Cartesian coordinates

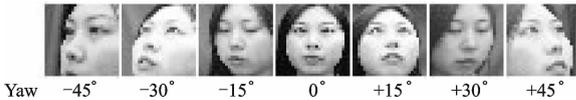


图4 区分Yaw角度的分类

Fig. 4 Classification with distinguishing the Yaw angle

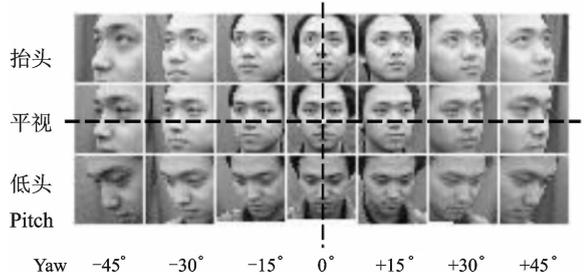


图5 区分Yaw角度和Pitch角度的分类

Fig. 5 Classification with distinguishing Yaw angle and Pitch angle

### 3.2 网络结构与配置

本文研发的人脸姿态分类检测系统由3层RBM网络与Softmax分类器叠加组成,如图6所示。由多层RBM和Softmax分类器组成的网络有时也称为Deep Belief Network。

对于深度网络系统其结构配置的主要参数是层数及各层的单元数。系统中每一层的分析处理都可以理解为系统对输入特征作了一次变换。因此增加层数将有利于揭示原始信息中更高层的内涵,但过多的层数不但增加了处理计算量,也可能会引起信息损失;在既定的层中单元数目过多可能导致出现过配(Overfit),单元数目过少又可能导致信息损失。合理选择网络层数和各层的单元数,是研发深度学习系统的一项重要工作内容。本文研发的系统,对于3种类别的检测,都采用相同的配置。输入层的单元数为2 352(当只采用灰度特征时为784),第1层RBM的隐藏层的单元数为1 000,第2层和第3层RBM的隐藏层单元数分别为500和600。分类器输出层单元数分别与分类类别数目相对应,即当分类类别分别为3,7,21时,其输出单元数目也是3,7,21。本文系

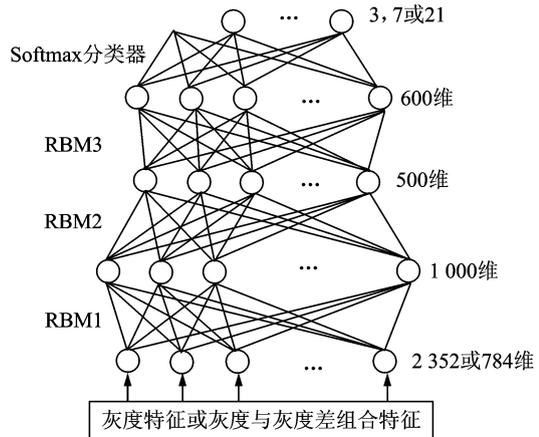


图6 网络结构示意图

Fig. 6 Diagram of network

统所选的 RBM 层数与各层单元数,在测试实践中经过有限次调整按经验选取,应当还有优化改进的空间。

### 3.3 网络训练

网络的训练分为两个阶段,即预训练和微调训练。在对网络预训练时,逐层对每层 RBM 网络独立地使用对比散度算法进行单层学习,即训练每层 RBM 的参数  $W, a, b$ , 其中参数  $W$  为网络提取特征的模板。图 7 为以灰度特征为输入的第 1 层 RBM 网络学习得到的 1 000 个特征中的 100 个特征模板。为了使图像输出效果更明显一些,本文在图像输出处理中人为增强了图像的对比度。从图 7 中可以明显看出,第 1 层 RBM 模型不但学习到了人脸器官的分布(如眼睛、鼻子和嘴巴等),而且也学习了人脸的轮廓描述。而人脸的姿态可以由人脸的轮廓以及器官的分布组合得到很好的估计,所以学习到的模板对姿态分类有利。

在微调训练时将训练好的 3 层 RBM 与 Softmax 分类器叠加起来,使用反向传播(Back propagation, BP)算法对网络进行整体统调,即根据分类结果反馈对预训练的各层 RBM 参数  $W, a, b$  及分类器层网络参数进行调整,使分类正确度达到最大。本文研发的系统对 3 种类型检测虽然采用了相同的结构和配置,但用于测试不同类型的检测系统,经过不同训练后的网络参数不同。

## 4 实验测试结果及分析

本文实验样本选自 CAS-PEAL-R1<sup>[20]</sup> 人脸数据库和 OFD(Oriental face database)<sup>[21]</sup> 人脸数据库。CAS-PEAL-R1 人脸库是 CAS-PEAL 人脸库的一个公开子集,在其姿态图像子库中每人有 21 种不同姿态的人脸图像,其姿态变化包括 7 种 Yaw 角度变化,每种 Yaw 姿态又包含 3 种 Pitch 角度变化(抬头、平视和低头)。在实验中,选取 939 人的 Yaw 角度为  $0^\circ, \pm 15^\circ, \pm 30^\circ$  及  $\pm 45^\circ$  的 21 类图像做样本。OFD 人脸数据库的视点子库中其 Pitch 姿态均为平视, Yaw 角度从  $-90^\circ \sim +90^\circ$ , 分类间隔为  $10^\circ$ 。将 OFD 人脸库中的 Yaw 角度为  $0^\circ, \pm 30^\circ$  图像提取出来,合并到 CAS-PEAL-R1 中的平视类图像中组成实验样本所用图像。

### 4.1 本文方法的分类实验结果

实验中将 CAS-PEAL-R1 人脸库中人脸姿态图像按照 Yaw(偏航)角度和 Pitch(仰俯)角度不同,每类随机取 800 张图像做训练集,剩余图像做测试集。并在 OFD 人脸库中选取 Yaw 角度为  $0^\circ, \pm 30^\circ$  3 类图像,每类 800 张合并到训练集中,以增加训练集样本数目。分别以图像灰度特征、灰度与灰度差组合特征为输入,输入到网络深层网络中,测试结果如表 1 所示。本次实验使用的计算机 CPU 配置为 Intel(R) Core(TM) i5 @ 3.10 GHz 3.5 GHz 4 核, 8 GB 主存,软件实验平台为 Matlab 7.11.0(R2010b), Win7 64 位操作系统。测试时间为将单张样本输入训练好的深层系统模型并得出估计姿态类所需的平均测试时间。

从表 1 中可以看出,无论是对 Yaw 角度、Pitch 角度或 Yaw 及 Pitch 角度组合分类的检测,其分类类型的正确识别率均高于 95%。与以单一灰度特征为输入相比,使用以灰度与灰度差组合特征为输入的网络取得了更好的识别效果。

### 4.2 实验结果与相关文献的对比

文献[10,11]给出了在 CAS-PEAL 人脸数据库上关于人脸姿态估计的新测试算法及测试结果,与本文提出方法进行对比,结果如表 2 所示。从表 2 中可以看出,本文提出方法明显优于文献[10,11]提出的算法。

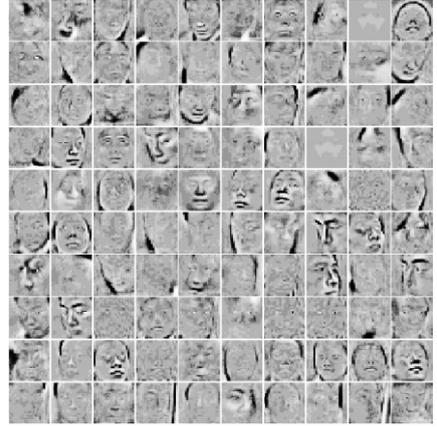


图 7 单层 RBM 网络学习所得的部分特征  
Fig. 7 Subset of the features learned by single layer of RBM

表 1 本文算法对不同姿态分类方法的识别率

Tab. 1 Recognition rate of different pose classification method with proposed algorithm

输入特征	性能	7 类	3 类	21 类
		Yaw	Pitch	Yaw&Pitch
灰度特征	识别率/%	97.10	97.43	95.23
	测试时间/ms	12.0	11.9	13.1
灰度差与灰度组合特征	识别率/%	97.87	98.05	95.74
	测试时间/ms	14.3	14.1	14.5

表 2 本文算法与最新的相关算法的识别率对比

Tab. 2 Comparison of recognition rate of proposed algorithm and latest algorithms %

分类类别	文献[10]	文献[11]	本文所提方法
7 类(Yaw)	86.85	93.29	97.87
21 类(Yaw&Pitch)		88.24	95.74

## 5 结束语

深度学习技术具有从大数据训练集中自动提炼反映研究对象复杂内涵特征的能力。通过本文中第 1 层 RBM 的特征模板可视化显示和与其他浅层学习算法对姿态分类结果的对比,验证了深度学习的效能。本文构建的深度学习人脸姿态检测分类实验系统在 CAS-PEAL-R1 公开人脸库上的实验测试,无论是对单独的俯仰 Pitch 角度、偏航 Yaw 角度或它们的组合分类,实验系统均获得了很高的正确率(识别率均高于 95%)。现有文献发表的在机器视觉及图像识别成功应用的深度学习系统,其大多只取图像的灰度特征为输入数据。在本文实验系统中,采用灰度与灰度差组合特征作为输入数据,取得的分类效果优于只取灰度值输入的分类效果。测试结果表明,在输入数据中增加与学习目标相关度高的信息内容,特别是含有互补性非冗余的信息,有助于深度学习提高性能。

## 参考文献:

- [1] 韩兆翠. 基于多线索的人脸识别认证[D]. 大连:大连理工大学,2013.  
Han Zhaocui. Face recognition with multiple cues[D]. Dalian:Dalian University of Technology,2013.
- [2] Murphy-Chutorian E,Trivedi M M. Head pose estimation in computer vision;A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2009,31(4):607-626.
- [3] Shafi M,Chung P W H. Face pose estimation from eyes and mouth[J]. International Journal of Advanced Mechatronic Systems,2010,2(1):132-138.
- [4] 梁国远,查红彬,刘宏. 基于三维模型和仿射对应原理的人脸姿态估计方法[J]. 计算机学报,2005,28(5):792-799.  
Liang Guoyuan,Zha Hongbin,Liu Hong. Face pose estimation based on 3D model and affine correspondences[J]. Journal of Computers,2005,28(5):792-799.
- [5] McKenna S J,Gong S. Real-time face pose estimation[J]. Real-Time Imaging,1998,4(5):333-347.
- [6] Ng J,Gong S. Composite support vector machines for detection of faces across views and pose estimation[J]. Image and Vision Computing,2002,20(5):359-368.
- [7] Krüger N,Pötzsch M,vonderMalsburg C. Determination of face position and pose with a learned representation based on labelled graphs[J]. Image and Vision Computing,1997,15(8):665-673.
- [8] Li S Z,Lu X G,Hou X W, et al. Learning multiview face subspaces and facial pose estimation using independent component analysis[J]. IEEE Transactions on Image Processing,2005,14(6):705-712.
- [9] Ma Bingpeng,Zhang Wenchao,Shan Shiguang, et al. Robust head pose estimation using LGBP[C]//18th International Conference of IEEE on Pattern Recognition. Hong Kong,China,IEEE,2006:512-515.

- [10] Shafi M, Iqbal F, Ali I. Face pose estimation using distance transform and normalized cross-correlation[C]// International Conference of IEEE on Signal and Images Processing Applications. Kuala Lumpur, IEEE, 2011; 186-191.
- [11] Han Zhaoqui, Su Tieming, Tang Xusheng, et al. Face recognition with integrating multiple cues[J]. *Journal of Signal Processing Systems*, 2014, 74(3): 391-404.
- [12] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [13] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1798-1828.
- [14] Hinton G, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [15] Tang Y, Salakhutdinov R, Hinton G. Robust Boltzmann machines for recognition and denoising[C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI: IEEE, 2011; 186-191.
- [16] 戴礼荣, 张仕亮. 深度语音信号与信息处理: 研究进展与展望[J]. *数据采集与处理*, 2014, 29(2): 171-179.  
Dai Lirong, Zhang Shiliang. Deep speech signal and information processing: Research progress and prospect [J]. *Journal of Data Acquisition and Processing*, 2014, 29(2): 171-179.
- [17] 杨阳, 张文生. 基于深度学习的图像自动标注算法[J]. *数据采集与处理*, 2015, 30(1): 88-98.  
Yang Yang, Zhang Wensheng. Image auto-annotation based on deep learning[J]. *Journal of Data Acquisition and Processing*, 2015, 30(1): 88-98.
- [18] Hinton G E. Training products of experts by minimizing contrastive divergence[J]. *Neural Computation*, 2002, 14(8): 1771-1800.
- [19] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Charles Rich: IEEE, 2001; 511-518.
- [20] Gao Wen, Cao Bo, Shan Shiguang, et al. The CAS-PEAL large-scale chinese face database and baseline evaluations[J]. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 2008, 38(1): 149-161.
- [21] 西安交通大学. OFD 东方人脸数据库[EB/OL]. <http://www.aiar.xjtu.edu.cn/group/face/Chinese/Homepage.htm>, 2007-04-08/2012-08-19.  
Xi'an Jiao Tong University. OFD oriental face database[EB/OL]. <http://www.aiar.xjtu.edu.cn/group/face/Chinese/Homepage.htm>, 2007-04-08/2012-08-19.

#### 作者简介:



**苏铁明** (1972-), 男, 博士, 讲师, 研究方向: 计算机视觉、模式识别及 CAD/CAM, E-mail: tiemings@dlut.edu.cn.



**程福运** (1989-), 男, 硕士研究生, 研究方向: 机器学习、模式识别。



**韩兆翠** (1983-), 女, 博士, 研究方向: 计算机视觉、模式识别。

**欧宗瑛** (1936-), 男, 教授, 研究方向: 计算机辅助设计、人工智能及应用。

