

核主成分分析网络的人脸识别方法*

胡伟鹏, 胡海峰, 顾建权, 李昊曦
(中山大学电子与信息工程学院, 广东 广州 510006)

摘要: 主成分分析网络 (principal component analysis network, PCANet) 是当前流行深度学习模型, 在人脸识别、目标识别、纹理分类和手写体数字识别等方面有广泛应用。在 PCANet 方法基础上, 该文提出基于核主成分分析网络 (kernel principal component analysis network, KPCANet) 的人脸识别方法。算法由四部分组成: 主成分分析 (PCA)、核主成分分析 (KPCA)、二值化哈希、分块直方图。在 Extended Yale B 和 AR 两个经典人脸库上的实验结果表明, 所提方法在识别性能上优于 PCANet 方法, 算法对于不同光照、表情变化下的人脸有更好的识别率和鲁棒性。

关键词: 核主成分分析网络; 深度学习; 人脸识别; 核变换

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 0529-6579 (2016) 05-0048-05

Kernel principal component analysis network method for face recognition

HU Weipeng, HU Haifeng, GU Jianquan, LI Haoxi

(School of Electronics and Information Engineering, Sun Yat-sen University,
Guangzhou 510006, China)

Abstract: Principal component analysis network (PCANet) is a popular deep learning classification method, which has caused wide attention in the area of computer vision due to its practical applications in face recognition, hand-written digit recognition, texture classification, and object recognitions. On the basis of PCANet. The kernel principal component analysis network (KPCANet) method is proposed for face recognition. The model is constructed by four processing components, including principal component analysis (PCA), kernel principal component analysis (KPCA), binary hashing, and block-wise histograms. The performance of the proposed method is evaluated using two public face datasets, i. e., Extended Yale B database and AR face database. The results show that KPCANet outperforms PCANet method. Especially when the face images have large variations about illuminations and expressions, KPCANet gives better recognition results.

Key words: kernel principal component analysis network; deep learning; face recognition; kernel transformation

基于视觉语义的人脸图像分类一直是一个充满挑战的研究领域, 不仅待识别的人脸图像种类多, 在每一类内部也存在较大差别, 包括光照变化、对

齐因素、形变因素、遮挡因素等等。针对这些问题, 学者们做了各种努力, 提出不同的模型方法来应对。当前, 人脸识别方法大致分为三类: 基于几

* 收稿日期: 2016-01-28

基金项目: 国家自然科学基金资助项目 (60802069, 61273270); 广东省自然科学基金资助项目 (2014A030313173)

作者简介: 胡伟鹏 (1993年生), 男; 研究方向: 计算机视觉, 模式识别; 通讯作者: 胡海峰; E-mail: huhaif@mail.sysu.edu.cn

何特征的方法、基于模板的方法（特征脸、线性判别分析、神经网络等）和基于模型的方法^[1-3]。深度学习作为机器学习的一个研究领域近几年受到广泛关注。该方法模拟大脑的深度组织结构，通过组合低层特征获取更抽象、更有效的高层语义信息。许多学者对深度学习进行了广泛的研究，提出各种不同架构的深度学习。Lee 等^[4]提出卷积深度置信网络（convolutional deep belief networks, CDBN），能够提取图像高级视觉特征。Christian 等^[5]提出深度神经网络（deep neural networks, DNNs），能提取图像几何信息达到更好分类效果，但存在过拟合和运算量大等缺点。Christian 等^[6]又提出卷积神经网络（convolutional neural networks, CNNs），存在网络参数少和简化训练的优点。

主成分分析网络（principal component analysis network, PCANet）在人脸识别、目标识别、纹理分类和手写体数字识别等方面有广泛应用^[7]，该模型的提出对已有传统深度学习模型（例如卷积神经网络结构以及小波分散式网络结构）构成一定的挑战^[8-9]。值得指出的是，传统 PCANet 方法的两层 PCA 变换只起扩大感受视野作用，PCA 层非线性拟合能力较差。针对这些问题，我们在 PCANet 的基础上，通过引入核（kernel）变换^[10-11]，将数据映射到线性可分高维空间，进行降维操作并获取滤波器，提高模型的非线性拟合能力。本文提出的新型核主成分分析网络（kernel principal component analysis network, KPCANet）方法由 4 部分组成：主成分分析（PCA），核主成分分析（kernel principal component analysis, KPCA），二值化哈希，分块直方图。实验表明 KPCANet 对于变化的光照和不同表情人脸图像有更佳识别效果，显示了更好的鲁棒性。

1 核主成分分析网络方法

PCANet 深度网络是对当前流行的卷积神经网络（convolutional neural networks, CNNs）与哈希直方图方法的一个结合^[6-12]。其 PCA 层对应 CNNs 网络中的卷积层，起到了神经元局部感受野和权值共享的作用，哈希层则起到了非线性拟合与降维功能，对应 CNNs 中的激活函数与池化层。进行两层 PCA 层主要能起到扩大感受视野的作用，PCA 层只进行线性变换，非线性拟合能力较差，不能保证原数据集是线性可分的。针对这一问题我们将原网络结构中的第二层 PCA 层换成 KPCA 层，数据在第一层 PCA 卷积层输出以后将被非线性映射到高维核空间中进行 PCA 降维，该过程与卷积网络中对卷积层线性变化后经过激活函数处理的操作有类似的思想^[6]，整个 KPCA 网络进行两层卷积 + 激活函数层，进行降维与分类（算法流程如图 1 所示）。这一处理相比 PCANet，能提高模型的非线性拟合能力，达到提升分类精度的效果。

设定输入的训练图像数目为 N ，即 $\{I_i\}_{i=1}^N$ ，每幅输入图像的大小为 $m \times n$ ，假定所有层中块（patch）大小为 $k_1 \times k_2$ 。PCANet 的 PCA 滤波器通过输入的训练样本 $\{I_i\}_{i=1}^N$ 获得。

1.1 第一层变换

对输入的图像 I_i ，每个像素点取大小为 $k_1 \times k_2$ 块，经过 patch 向量化将所有点信息有序的排列起来，记为 $x_{i,1}, x_{i,2}, \dots, x_{i,mn} \in \mathbf{R}^{k_1 k_2}$ ，其中 $x_{i,j}$ 表示 j_{th} 向量化的 patch。用每个向量化 patch 减去其均值，得到 $\bar{X}_i = [\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,mn}]$ ， $\bar{x}_{i,j}$ 表示均值化后的向量。将所有的训练样本构建相同的矩阵，得到：

$$X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in \mathbf{R}^{k_1 k_2 \times Nmn} \quad (1)$$

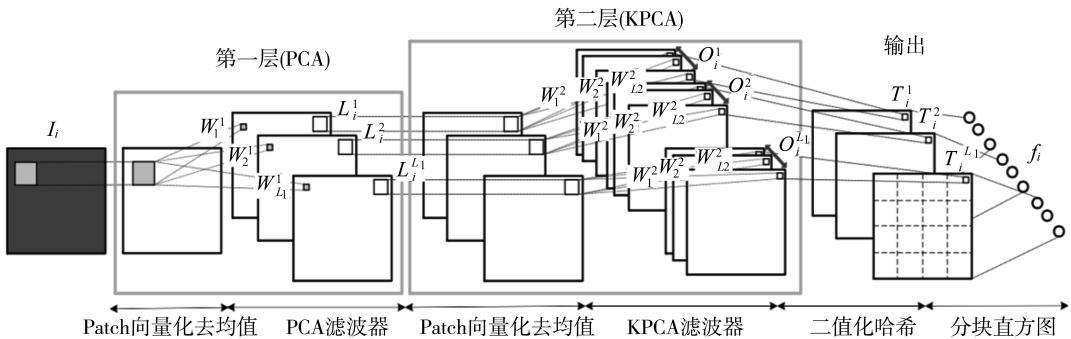


图 1 核主成分分析网络（两层结构）算法流程图

Fig. 1 Flowchart of our proposed kernel principal component analysis network (two-stage)

假定第 i 层的滤波器数目为 L_i , PCA 最小化重构误差满足下式:

$$\begin{aligned} \min_{v \in \mathbf{R}^{k_1 k_2 \times L_1}} \|X - VV^T X\|_F^2 \\ \text{s. t. } V^T V = I_{L_1} \end{aligned} \quad (2)$$

式中, I_{L_1} 大小为 $L_1 \times L_1$ 的单位矩阵。 XX^T 的 L_1 个主特征向量为上式的解, PCA 滤波器为:

$$w_l^1 = \text{mat}(q_l(XX^T)) \in \mathbf{R}^{k_1 \times k_2} \quad (3)$$

上式将向量 $v \in \mathbf{R}^{k_1 k_2}$ 映射为 $w \in \mathbf{R}^{k_1 \times k_2}$ 的函数, $q_l(XX^T)$ 表示 XX^T 的第 l 个主特征向量^[13], $l = 1, 2, \dots, L_1$ 。较大靠前的特征值对应的特征向量获得了所有均值训练块的主要变化。

1.2 第二层变换

设定第一层第 l 个 PCA 滤波器输出:

$$I_i^l = I_i * w_l^1, i = 1, 2, \dots, N \quad (4)$$

I_i 与 w_l^1 作卷积运算, 图像 I_i 与滤波器 w_l^1 作卷积运算前, 图像的边界用零值填充。与第一层相同, 通过对 I_i^l 像素取块、向量化、去均值, 得到 $\bar{Y}_i^l = [\bar{y}_{i,1}, \bar{y}_{i,2}, \dots, \bar{y}_{i,mn}] \in \mathbf{R}^{k_1 k_2 \times mn}$, \bar{Y}_i^l 矩阵的行向量 $v_i \in \mathbf{R}^{mn}$ 通过滑动窗口 patch 作用于“图像” I_i^l 获得, \bar{Y}_i^l 矩阵的行数等于 patch 矩阵的元素个数 $k_1 k_2$ 。现将 \bar{Y}_i^l 中的行向量 v_i 看作样本, 进行核变换。常用的核函数有多项式核函数^[9-10]、高斯核函数、多层感知核函数和线性核函数等, 其定义分别如下:

1) Gaussian 核函数:

$$\mathbf{K}(v_i, v_j) = \exp\left(-\frac{\|v_i - v_j\|^2}{2\sigma^2}\right) \in \mathbf{R}^{k_1 k_2 \times k_1 k_2} \quad (5)$$

式中 $v_i, v_j \in \mathbf{R}^{mn}$ 表示矩阵 \bar{Y}_i^l 中的行向量, σ 是实参数, $i, j = 1, 2, \dots, k_1 k_2$ 。

2) PolyPlus 核函数:

$$\mathbf{K}(v_i, v_j) = [(v_i v_j^T) + 1]^d \in \mathbf{R}^{k_1 k_2 \times k_1 k_2} \quad (6)$$

式中 $v_i, v_j \in \mathbf{R}^{mn}$ 表示矩阵 \bar{Y}_i^l 中的行向量, d 是实参数。

3) Polynomial 核函数:

$$\mathbf{K}(v_i, v_j) = (v_i v_j^T)^d \in \mathbf{R}^{k_1 k_2 \times k_1 k_2} \quad (7)$$

式中 $v_i, v_j \in \mathbf{R}^{mn}$ 表示矩阵 \bar{Y}_i^l 中的行向量, d 是实参数。 \bar{Y}_i^l 经过核变换后记为 $\mathbf{K}_i \in \mathbf{R}^{k_1 k_2 \times k_1 k_2}$, 将 $\{\bar{Y}_1^l, \bar{Y}_2^l, \dots, \bar{Y}_N^l\}$ 进行核变换后, 串联成矩阵 \mathbf{K} , 也就是:

$$\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2, \dots, \mathbf{K}_N] \in \mathbf{R}^{k_1 k_2 \times N k_1 k_2} \quad (8)$$

再求特征向量, 并取前 l 个特征向量, 即:

$$w_l^2 = \text{mat}(q_l(KK^T)) \in \mathbf{R}^{k_1 \times k_2} \quad (9)$$

每个输入 I_i^l , 对应 L_2 个输出, I_i^l 与 w_l^2 作卷积运算:

$$O_i^l = \{I_i^l * w_l^2\}_{l=1}^{L_2}, l = 1, 2, \dots, L_2 \quad (10)$$

共 $L_1 L_2$ 个输出。重复上面步骤, 可以构建多层网络结构。

1.3 输出层: 哈希和直方图

第一层中 I_i^l 的 L_1 个并行输入, 其中每一个输入在第二层中对应 L_2 实值输出 $\{I_i^l * w_l^2\}_{l=1}^{L_2}$ 。二值化输出 $\{H(I_i^l * w_l^2)\}_{l=1}^{L_2}$, $H(\cdot)$ 是单位阶跃函数。对每个像素 L_2 个二值化向量加权输出为十进制整数。 O_i^l 中 L_2 个输出转化成值为整数“图像”。

$$T_i^l = \sum_{l=1}^{L_2} 2^{l-1} H(I_i^l * w_l^2) \quad (11)$$

像素值的范围为 $[0, 2^{L_2} - 1]$, 对于 L_1 中的每个图像 T_i^l 划分成 B 个 block。对每个 block 进行直方图统计^[14], 将 B 个直方图组合成一个向量, 表示为 $\text{Bhist}(T_i^l)$ 。通过编码处理后, 输入图像 I_i 的特征输出为

$$f_i = [\text{Bhist}(T_i^1), \dots, \text{Bhist}(T_i^{L_1})]^T \quad (12)$$

式中, $f_i \in \mathbf{R}^{(2^{L_2})L_1 B}$ 。KPCANet 模型参数: 每层滤波器数目 L_1, L_2 , 模型层数, 输出层中局部直方图块大小, 块重叠区域比率。

2 实验结果与分析

我们将所提的 KPCANet 算法在 Extended Yale B 和 AR 两个经典人脸库进行了实验验证。

2.1 Extended Yale B 数据库

在 Extend Yale B 人脸数据库的实验中, 共 38 个人, 每人 64 张人脸图像 (2 432 张人脸图, 图像尺寸为 100×100), 这些图像是不同光照条件下拍摄的正面人脸图像。图 2 是 Extend Yale B 数据库中同一个人不同光照条件下的 20 张正面人脸图像。实验时, 我们对每个人随机抽取 5 张人脸图像进行训练, 共 190 个训练样本, 其余的作为测试, 测试样本数量为 2 242 个, 进行 20 次独立实验, 取均值作为实验结果。



图 2 Extended Yale B 不同光照条件正面人脸
Fig. 2 Examples frontal face images with various illumination conditions in Extended Yale B database

表 1 Extend Yale B 库中人脸识别率

Table 1 Recognition accuracy in Extend Yale B database

算法	识别率/%
PCANet	98.49
KPCANet-Pn	98.79
KPCANet-Pp	98.80
KPCANet-Gs	98.82

为了比较分析, PCANet 和 KPCANet 设置相同模型的参数。即: 2 层模型结构, 每层滤波器数目均设置 8, 输出层中局部直方图块 (block) 大小 7×7 , 块重叠区域比率 0.5。另外, KPCANet 针对不同核函数的参数赋值如下: Gaussian 核参数 $\sigma = 6$, PolyPlus 核参数 $d = 2$, Polynomial 核参数 $d = 6$ 。对于不同核函数, 我们用 KPCANet-Gs、KPCANet-Pp 和 KPCANet-Pn 分别表示。在 Extended Yale B 库中实验结果如表 1 所示。由表可得, PCANet 的识别率为 98.49%, KPCANet 在不同核函数下算法识别率均高于 PCANet, 其中高斯核函数的性能最好, 识别率为 98.82%。高斯核取不同 σ 值时的识别性能如图 3 所示, 实验表明, 高斯核取不同 σ 值对人脸识别率有一定影响, σ 为 6 的时人脸识别率最高。

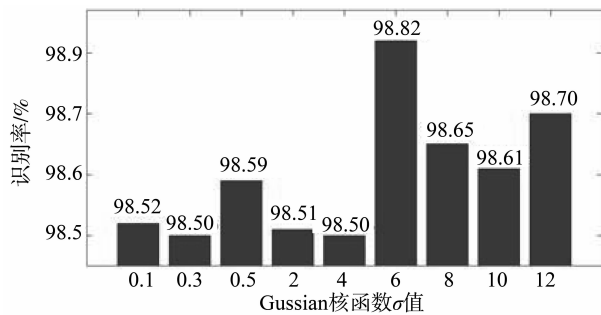


图 3 Gaussian 核取值 σ 与人脸识别率

Fig. 3 Gaussian kernel value and recognition accuracy

2.2 AR 数据库实验

AR 数据库包含 50 个人, 每人 14 张照片 (共 700 张人脸图像, 图像大小为 165×120), 这些图像均为在光照变化不大的条件下拍摄的正面人脸图像, 每个人包含不同表情 (如微笑、张嘴、闭眼等)。图 4 是 AR 库同一个人的 14 张图片。实验中每个人随机抽取 7 张人脸图像作为训练, 共 350 个训练样本, 其余的作为测试, 测试样本数量为 350 个, 经过 20 次独立实验, 取均值作为实验结果。

实验结果如表 2 所示, 我们发现 PCANet 的识别率为 99.59%, KPCANet 不同的核函数下算法识

别率均高于 PCANet, 其中用高斯核函数性能最好, 识别率为 99.76%。



图 4 AR 库不同表情正面人脸

Fig. 4 Examples frontal face images with various expression in AR database

表 2 AR 库中人脸识别率

Table 2 Recognition accuracy in AR database

算法	识别率/%
PCANet	99.59
KPCANet-Pp	99.62
KPCANet-Pn	99.73
KPCANet-Gs	99.76

2.3 结果分析

通过对 Extended Yale B 和 AR 两个人脸数据库中的实验表明, 所提的 KPCANet 模型比 PCANet 方法在识别性能上有一定的提高。Extended Yale B 人脸数据库中, KPCANet 识别率高达 98.82%, AR 人脸数据库中, KPCANet 识别率高达 99.76%, 结果表明我们的方法对光照、表情变化鲁棒, 有很好的适应特性。

3 结语

本文提出一种新型的核主成分分析网络方法用于人脸识别, 我们的模型包括对输入图像进行主成分分析、核主成分分析、二值化哈希和分块直方图分析等。与传统方法相比, 本算法引入“核”变换, 提高了模型的非线性拟合能力, 达到提升分类精度的效果。实验结果表明所提方法在识别性能上优于 PCANet 方法, 对于不同光照、表情下的人脸变化具有更好的识别率和鲁棒性。

参考文献:

[1] TURK M, PENTLAND A. Face recognition using eigenfaces [C] // IEEE Conference on Computer Vision and Pattern Recognition, 1991: 586 - 590.
 [2] FISHER R. The use of multiple measurements in taxonomic problems [J]. Annals of Eugenics, 1936, 7(2): 179 - 188.

- [2] 郭昌. 小波变换与 HMT 模型的图像插值算法[J]. 中山大学学报(自然科学版), 2012, 51(3): 55 - 59.
- [3] VENKATESHWAR R M, BHAGYA R V. Image resolution enhancement technique using lifting wavelet and discrete wavelet transforms [C] // Proceeding of the 3rd International Conference on Innovations in Computer Science and Engineering, 2015, 413: 235 - 239.
- [4] CHRYSAFIS C, ORTEGA A. Line - based, reduced memory, wavelet image compression [J]. IEEE Transactions on Image Processing, 2000, 9(3): 378 - 389.
- [5] 刘正光, 申旭刚. 基于行的小波变换及其在图像压缩中的应用[J]. 中国图像图形学报, 2004, 9(3): 370 - 374.
- [6] 张宏伟, 刘正光, 陈红新. 应用提升格式实现的基于行小波变换的图像压缩算法[J]. 计算机应用, 2005, 25(7): 1626 - 1628.
- [7] YE L, GUO J, NUTTER B, et al. Low - memory - usage image coding with line - based wavelet transform [J]. Optical Engineering, 2011, 50(2): 1 - 11.
- [8] ERISTI H. Fault diagnosis system for series compensated transmission line based on wavelet transform and adaptive neuro - fuzzy inference system [J]. Measurement, 2013, 46(1): 393 - 401.
- [9] 杨汉生, 孔鲲鹏, 许磊. 基于正交多项式的自适应提升格式[J]. 南京理工大学学报(自然科学版), 2009, 33(3): 307 - 311.
- [10] SHOBA P, MARIA L L, MOHAN V, et al. Landsat image compression using lifting scheme [C] // Proceeding of International Conference on Communications and Signal Processing (ICCSPP). Tamilnadu Melmaruvathur, India, 2014: 1963 - 1968.
- [11] 高东, 张贝克, 吴重光. 基于提升格式的过程数据在线去噪方法及其应用[J]. 计算机应用研究, 2008, 25(10): 3198 - 3200.
- [12] ZHANG L, QIU B. Edge - preserving image compression using adaptive lifting wavelet transform [J]. International Journal of Electronic, 2015, 102(7): 1190 - 1203.
- [13] GUO F, LIU X, ZHAO C. The image enhancement technology based on lifting wavelet [J]. Applied Mechanics and Materials, 2014, 513/514/515/516/517: 3261 - 3264.
- [14] ZENG W, DALY S, LEI S. An overview of the visual optimization tools in JPEG 2000 [J]. Signal Processing; Image Communication, 2002, 17(1): 85 - 104.
- [15] 丁晓峰, 何凯霖. 基于改进提升小波变换 SPIHT 的图像压缩算法[J]. 计算机测量与控制, 2014, 22(11): 3670 - 3672.

(上接第 51 页)

- [3] 左军, 周灵, 孙亚民. 基于 RBF 神经网络 PCA 变换的识别技术 [J]. 中山大学学报(自然科学版), 2014, 53(6): 135 - 140.
- [4] LEE H, GROSSE R. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations [C] // International Conference on Machine Learning, 2009: 609 - 616.
- [5] CHRISTIAN S, TOSHEV A, ERHAN D. Deep neural networks for object detection [C] // Advances in Neural Information Processing Systems, 2013, 2553 - 2561.
- [6] CHRISTIAN S, LIU W, JIA Y Q. Going deeper with convolutions [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1 - 9.
- [7] CHAN T H, JIA K, GAO S H, et al. Pcanet: A simple deep learning baseline for image classification [J]. TIP, 2015, 24(12): 5017 - 5032.
- [8] BRUNA J, MALLAT S. Invariant scattering convolution net-works [J]. TPAMI, 2013, 35(8): 1872 - 1886.
- [9] SIFRE L, MALLAT S. Rotation, scaling and deformation invariant scattering for texture discrimination [J]. CVPR, 2013, 9(4): 1233 - 1240.
- [10] HOFMANN T, SCHOLKOPF B, ALEXANDER J S. Kernel methods in machine learning [J]. Annals of Statistics, 2008, 36(3): 1171 - 1220.
- [11] 黄啸. 支持向量机核函数的研究 [D]. 苏州: 苏州大学, 2008.
- [12] ALEX K. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25(2): 195 - 203.
- [13] 黄晓生, 严浩, 曹义亲, 等. 基于小波高频奇异值分解的无参考模糊图像质量评价 [J]. 中山大学学报(自然科学版), 2014, 53(6): 165 - 171.
- [14] 赵慧民, 蔡君, 魏文国. 一种分块压缩感知变采样率的指纹图像水印算法 [J]. 中山大学学报(自然科学版), 2014, 53(5): 14 - 20.