

文章编号: 1006-2475(2018) 11-0046-05

# 一种结合 Gabor 小波与深度学习的人脸识别方法

潘峥嵘, 王 震

(兰州理工大学电气工程与信息工程学院, 甘肃 兰州 730050)

**摘要:** 针对人脸识别中识别效果易受光照、姿态等因素影响和浅层学习方法不能有效提取人脸图像抽象特征的问题, 提出一种结合 Gabor 小波与深度学习的人脸识别方法。该方法首先利用 Gabor 小波变换获取不同尺度和方向的人脸 Gabor 特征, 通过下采样和受限玻尔兹曼机(RBM)对 Gabor 特征进行有效降维; 其次将降维后的特征作为深度信念网络(DBN)的输入, 并使用对比散度算法训练 DBN; 最后利用标签数据对 DBN 进行有监督微调, 网络顶层附加 Softmax 分类器对提取后的特征进行分类。所提方法在 ORL、UMIST 和 Yale-B 人脸库上的识别率分别达到了 98.72%、96.51% 和 96.13%, 实验结果表明所提方法不仅识别效果明显优于其他现有方法, 而且对光照、姿态变化具有很好的鲁棒性。

**关键词:** Gabor 小波; 人脸识别; 深度学习; 深度信念网络; 受限玻尔兹曼机

中图分类号: TP391.4 文献标识码: A doi: 10.3969/j.issn.1006-2475.2018.11.009

## A Face Recognition Algorithm Based on Gabor Wavelet and Deep Learning

PAN Zheng-rong, WANG Zhen

(School of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** In order to reduce the negative effects of factors such as illumination and posture and solve the problem that shallow learning methods can't extract the abstract features of face images in face recognition, a face recognition algorithm based on the Gabor wavelet and deep learning was proposed. Firstly, the facial Gabor features of different scales and directions were obtained by Gabor wavelet transform, the dimensionality of Gabor features was reduced available by downsampling and Restricted Boltzmann Machine (RBM). Secondly, the features of dimensionality reduced were taken as the input of the Deep Belief Networks (DBN), and DBN was trained by the Contrastive Divergence algorithm. Finally, DBN was fine-tuned by labeled data. The Softmax classifier was used to classification for the features extracted, which was implemented at the top layer. The recognition rate reaches 98.72%, 96.51% and 96.13% respectively on ORL, UMIST and Yale-B face databases. The experiment results indicate that the proposed method is markedly better than other existing algorithms in recognition performance and achieves good robustness to changes in illumination and posture.

**Key words:** Gabor wavelet; face recognition; deep learning; DBN; RBM

## 0 引 言

人脸识别技术具有非侵犯性、可交互性强、方便快捷等优点, 一直是模式识别和人工智能领域中的一个研究热点, 广泛应用于移动支付、平安城市、刑事侦查等多个领域。

传统的特征提取方法主要有基于全局和基于局部的 2 种特征提取方法, 基于全局的特征提取方法主要有主成分分析 (Principal Component Analysis, PCA)、线性判别分析 (Linear Discriminant Analysis,

LDA) 等。Kirby 和 Shakhunaga 等人<sup>[1-2]</sup>提出的 Eigenface (PCA) 方法通过低维线性子空间对人脸进行表达, 识别率较低; Belhumeur 等人<sup>[3]</sup>提出的 Fisherface (LDA) 方法利用 Fisher 准则使类间和类内之间的散度比值最大化, 但对于小样本数据识别效果较差。这些基于全局的特征提取方法受光照、姿态等因素影响较大, 识别率明显低于基于局部的特征提取方法。基于局部的特征提取方法主要有 Gabor 小波、局部二值模式 (Local Binary Pattern, LBP) 等。Gabor 小波对图像边缘敏感, 具有良好的尺度选择和方向选择特性,

收稿日期: 2018-05-04

基金项目: 甘肃省自然科学基金资助计划项目 (1308RJZA273)

作者简介: 潘峥嵘 (1964-) 男, 山东蓬莱人, 兰州理工大学电气工程与信息工程学院教授, 学士, 研究方向: 机器人视觉、智能控制; 王震 (1990-) 男, 河南开封人, 硕士研究生, 研究方向: 图像处理、机器学习。

可以较好地描述图像纹理信息,并且对光照、姿态等变化具有一定的鲁棒性,是一种常用的人脸特征提取方法<sup>[4]</sup>。文献[5]提出一种结合 Gabor 小波变换与 BP 神经网络的人脸识别方法,但 BP 神经网络容易陷入局部最优<sup>[6]</sup>,识别效果并没有得到明显改善;文献[7]提出一种基于 Gabor 小波与 PCA 的方法进行人脸识别。文献[5]和文献[7]都属于浅层学习方法<sup>[8]</sup>,当给定有限数量的样本和计算单元时,浅层学习方法无法对复杂函数进行有效表达,文献[9]列出了一些浅层学习方法无法表达的函数,揭示了浅层学习方法的局限性。相比于浅层学习方法,深度学习可以实现对复杂函数任意精度的近似,并具有强大的特征提取能力<sup>[10-11]</sup>。文献[12]提出了一种 LBP 与深度信念网络(Deep Believe Network, DBN)相结合的人脸识别方法,取得了很好的识别效果,表明了深度学习优异的性能。

基于上述分析,本文提出一种结合 Gabor 小波与深度学习的人脸识别方法。该方法通过下采样和 RBM 对得到的 Gabor 特征进行有效降维,降维后的特征作为 DBN 的输入,自下而上,逐层学习更抽象的特征,并在最顶层利用 Softmax 分类器对提取到的抽象特征进行分类。实验结果验证了本文所提方法的有效性。

## 1 相关工作

### 1.1 人脸 Gabor 特征的提取

Gabor 小波是 Gabor 变换与小波理论的结合体,不但具有 Gabor 变换的空间局部性和方向性,同时也继承了小波变换的多分辨率特性。二维 Gabor 小波变换可以捕捉图像纹理特征,在图像处理领域应用广泛<sup>[13-14]</sup>。二维 Gabor 滤波器定义为:

$$\phi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right) \times [\exp(ik_{u,v}z) - \exp(-\frac{\sigma^2}{2})] \quad (1)$$

其中  $z(x, y)$  为图像像素点坐标;  $u$  代表 Gabor 滤波器的方向;  $v$  代表 Gabor 滤波器的尺度;  $k_{u,v}$  代表滤波器的中心频率,它描述 Gabor 滤波器不同方向和尺度上的响应,可将其表示为:

$$k_{u,v} = \begin{pmatrix} k_v \cos \varphi_u \\ k_v \sin \varphi_u \end{pmatrix} \quad (2)$$

其中  $k_v = k_{\max}/f^v$ ,  $k_v$  代表采样尺度;  $k_{\max}$  代表最大频率;  $\varphi_u = \pi u/8$ ,  $\varphi_u$  体现了滤波器方向选择性;  $\|k_{u,v}\|^2/\sigma^2$  用来补偿由频率决定的能量谱的衰减;  $\exp(-\|k_{u,v}\|^2 \|z\|^2/2\sigma^2)$  为 Gaussian 函数;  $\exp(ik_{u,v}z)$  为振荡函数,实部为余弦函数,虚部为正弦函数;  $\exp(-\sigma^2/2)$  为直流分量,用来降低图像灰度绝对值对滤波器的干扰,使滤波器对图像光照变化敏感性降低。

Gabor 特征图像通过 Gabor 滤波器来获取,当输入图像灰度值为  $M(z)$  时,此输入图像与 Gabor 滤波器的卷积定义为:

$$O_{u,v}(z) = M(z) * \phi_{u,v}(z) \quad (3)$$

其中  $O_{u,v}(z)$  代表卷积后的 Gabor 特征图像;  $*$  代表卷积运算。理论研究表明<sup>[15]</sup>,为了获取有效的 Gabor 特征,一般选用 5 个尺度和 8 个方向的 Gabor 滤波器组成滤波器组。Gabor 特征的相位信息会随着空间位置变化,而幅值特征的变化却相对稳定和平滑<sup>[16-17]</sup>。本文采用 Gabor 幅值特征。

取 ORL 数据库中的一个样本,将其与 5 个尺度和 8 个方向的 Gabor 滤波器实部组成的滤波器组进行卷积,令  $k_{\max} = \pi/2$ ,  $f = \sqrt{2}$ ,  $\sigma = 2\pi$ ,人脸 Gabor 幅值特征图像如图 1 所示。

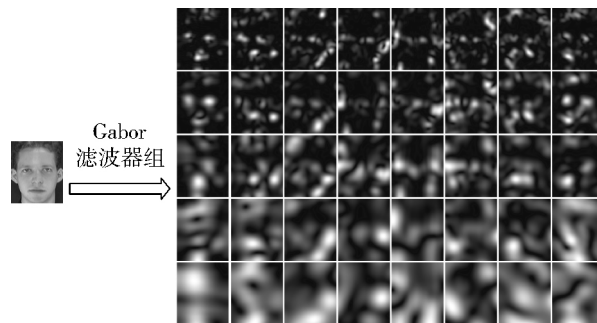


图1 人脸 Gabor 幅值特征图像

### 1.2 深度信念网络

DBN 是一种常用的深度学习模型, RBM 是组成 DBN 的基本结构单元。

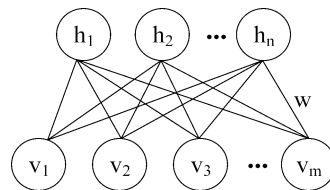


图2 RBM 结构

RBM 由可见层  $v$  和隐含层  $h$  组成,结构如图 2 所示。RBM 同层节点互不相连,层与层之间的节点互相连接,层内的每一个节点都是一个二值随机变量且满足玻尔兹曼分布。

RBM 是一种基于能量的概率模型,其能量函数被定义为:

$$E(v, h; \theta) = -\sum_{i=1}^m b_i v_i - \sum_{j=1}^n a_j h_j - \sum_{j=1}^n \sum_{i=1}^m W_{ji} v_i h_j \quad (4)$$

其中  $\theta = \{W, b, a\}$  是 RBM 模型的参数;  $W_{ji}$  是可见层节点  $i$  与隐藏层节点  $j$  之间的连接权值;  $b_i$  是可见层

节点  $i$  的偏置;  $a_j$  是隐藏层节点  $j$  的偏置;  $n$  为隐藏层节点个数。可见层与隐藏层的联合概率分布定义为:

$$p(v|h;\theta) = \frac{1}{Z(\theta)} \exp(-E(v|h;\theta)) \quad (5)$$

$$Z(\theta) = \sum_{v,h} \exp(-E(v,h;\theta)) \quad (6)$$

由以上公式可知,若已知可见层  $v$  的状态,隐藏层节点  $h_j$  的激活概率为:

$$p(h_j = 1 | v) = h(\sum_i W_{ji} v_i + b_j) \quad (7)$$

同样,已知隐藏层  $h$  的状态,可见层节点  $v_i$  的激活概率为:

$$p(v_i = 1 | h) = h(\sum_j W_{ji} h_j + a_i) \quad (8)$$

其中  $h(x)$  为 Sigmoid 激活函数。

由公式(7)和公式(8)可知,若 RBM 某层节点状态已知,可以得出下一层节点的激活概率,这是 RBM 的关键所在。

DBN 由多层 RBM 叠加而成,其典型结构如图 3 所示。DBN 的训练过程主要分为以下 2 步: 1) 输入向量作为可见层  $v$  的输入数据,利用对比散度(Contrastive Divergence, CD) 算法训练第一层 RBM,将训练完成的第一层 RBM 的输出作为第二层 RBM 的输入,以此类推,逐层训练每一个 RBM,完成对 DBN 的预训练。2) 使用标签数据并利用误差反向传播算法(Back Propagation, BP) 对网络进行微调,使 DBN 连接权值达到全局最优。

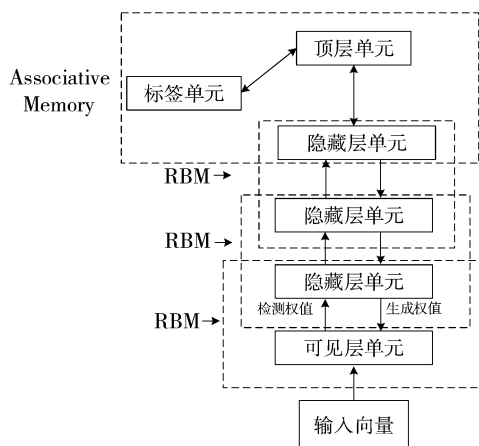


图3 DBN模型结构

## 2 Gabor 小波与 DBN 结合的人脸识别

### 2.1 基于下采样和 RBM 的特征降维

本文选用 5 个尺度和 8 个方向的滤波器实部组成滤波器组。假设输入样本  $M_i$  的分辨率为  $M \times N$ , 通过 Gabor 滤波器组变换后的特征维数则为  $M \times N \times 40$ 。若将其直接作为 DBN 的输入,不但冗余信息过多,而且会导致计算量大幅增加,并且若训练样本过少,模型很容易过拟合,故需要对 Gabor 特征进行降维。

本文提出一种基于下采样和 RBM 的降维方法,如图 4 所示。本文所提降维方法不但可以对 Gabor 特征进行有效降维,而且还可以提取其本质特征,这样更有助于分类。

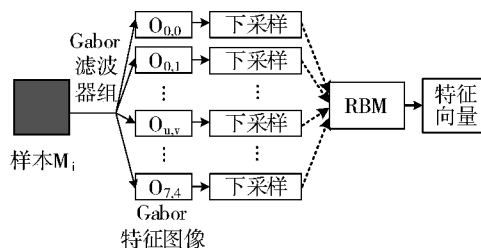


图4 基于下采样和 RBM 的特征降维

如图 4 所示,样本  $M_i$  通过 Gabor 滤波器组生成 40 幅 Gabor 特征图像  $O_{u,v}$ , 其中  $u = (0, 1, 2, 3, 4, 5, 6, 7)$ ,  $v = (0, 1, 2, 3, 4)$ 。先将获取的 Gabor 特征图像通过下采样初步对特征进行降维,具体方法是:使用一个  $2 \times 2$  的滑动窗口从左到右、从上到下依次在 Gabor 特征图像上进行步长为 2 的平均值采样,下采样后的特征维数降至原先的  $1/4$ 。但是,经过下采样后的特征维数依然很高,故后接一个 RBM 进行二次特征降维,具体方法是:将每幅 Gabor 特征图像下采样后的特征信息分别作为该 RBM 的输入,并使用 CD 算法训练该 RBM,进而对每幅 Gabor 特征图像下采样后的特征信息重新编码,实现特征降维的同时又提取了其本质特征,最终将提取到的所有特征信息组合并作为 DBN 的输入。

### 2.2 算法流程

本文算法将人脸 Gabor 特征与 DBN 相结合,具体流程如图 5 所示。

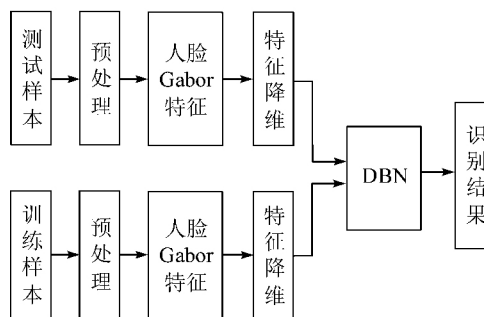


图5 算法流程图

算法具体步骤如下:

1) 使用双线性内插法将测试样本和训练样本降维至  $32 \times 32$ , 并进行直方图均衡化等预处理工作。

2) 将预处理后的测试样本和训练样本通过一个由 40 个不同尺度和方向的 Gabor 滤波器组成的 Gabor 滤波器组生成 40 幅人脸 Gabor 特征图像,通过下采样和训练完成的 RBM 对特征进行降维。图 4 中 RBM 结构为 40-5。

3) 将训练样本降维后的特征信息作为 DBN 可见层  $v$  的输入数据,并对 DBN 进行训练。首先,采用 CD 算法逐层训练每个 RBM,完成对 DBN 的预训练;其次利用标签数据并使用 BP 算法对整个网络进行有监督微调,使 DBN 连接权值达到全局最优。本文 DBN 结构为 1280-640-320-类别数,由 2 层 RBM 组成,第 1 层学习率为 0.005,迭代次数为 30;第 2 层学习率为 0.004,迭代次数为 30(在 ORL 人脸库上进行测试,当迭代 30 次时效果最佳);320-类别数这一层为 Softmax 回归层。微调时的学习率为 0.005,迭代次数为 600。

4) 当 DBN 训练完成以后,将测试样本降维后的特征作为 DBN 可见层  $v$  的输入,利用 DBN 进一步提取其抽象特征,在网络顶层附加一个 Softmax 分类器对测试样本进行分类。

### 3 实验结果与分析

本机硬件配置为 Intel Core™ i5-7500 CPU @ 3.40 GHz 8.00 GB 内存,软件环境为 Matlab R2016a。为了验证本文所提方法的有效性,选取文献[5]的 Gabor + BP 神经网络方法、文献[7]的 Gabor + PCA 方法、文献[18]的 PCA + SVM 方法和 DBN 方法进行对比实验,识别率取 10 次实验的平均值。

#### 3.1 ORL 人脸库实验

ORL 人脸库包含 40 个人共 400 幅图像,每人 10 幅在不同光照、姿态、表情下的图像,每幅图像分辨率为  $92 \times 112$ 。在 ORL 数据库中随机选取每人  $m$  ( $m = 3, 4, 5, 6, 7$ ) 幅图像作为训练样本,其余作为测试样本。对图 4 中 RBM 进行  $40 \times m$  轮循环训练,用得到的降维后的特征作为 DBN 的输入,完成对 DBN 的训练,DBN 结构为 1280-640-320-40。

由图 6 所示的实验结果可知,本文方法在不同规模训练样本下的识别率均最高。Gabor + PCA、Gabor + BP 神经网络等浅层学习方法虽然在一定程度上减弱了光照、姿态等因素对识别效果的影响,但对 Gabor 特征无法进行有效的后续处理,Gabor + PCA 方法识别率较低的原因在于较多的光照、姿态等因素累积到 PCA 的主向量上,影响了算法性能,而 Gabor + BP 神经网络方法识别效果较差的原因可能是 BP 神经网络陷入了局部最优。DBN 方法虽然能够学习到人脸图像的抽象特征,但 DBN 直接以像素级特征作为输入,无法学习到人脸图像的局部结构性特征,所学习到的特征对光照、姿态等变化较敏感,影响识别效果。当训练样本数为 280 时,本文方法识别率高达 98.72%,较 Gabor + BP 神经网络、Gabor + PCA 等浅

层学习方法分别高出 4.23% 和 6.25%,表明本文方法能够有效地提取人脸图像的抽象特征,较 DBN 方法高出 3.19%,表明本文方法能够降低光照、姿态等因素对识别效果的影响。

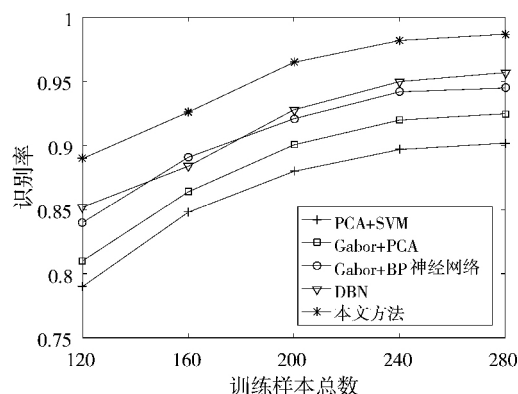


图6 ORL 数据库上识别率对比

#### 3.2 UMIST 人脸库实验

UMIST 人脸库包含 20 个人共 564 幅图像,包含从人脸侧面到正面的一系列姿态变化,每幅图像分辨率为  $220 \times 220$ 。在 UMIST 数据库中每人随机选取 19 幅图像,并在这 19 幅图像中随机选取  $m$  ( $m = 8, 10, 12, 14, 16$ ) 幅图像作为训练样本,其余作为测试样本。本实验采取与 ORL 人脸库上相同的实验策略进行参数配置。

由图 7 中的实验结果可知,由于 UMIST 人脸库中人脸姿态变化范围较大,各方法识别率较 ORL 人脸库均有所降低,PCA + SVM 方法下降最多,原因是经过 PCA 降维后得到的是人脸图像的全局特征,当受姿态等因素影响较大时,不能突出有利于分类的局部信息,并且 SVM 无法进一步提取其抽象特征,导致识别率大幅下降;而本文所提方法相比其他 4 种方法下降最少,识别率仍能达到 96.51%,更是比 PCA + SVM 方法高出 11.27%,表明本文方法对姿态这一非线性变化具有一定的鲁棒性,并且具有很强的特征学习能力。

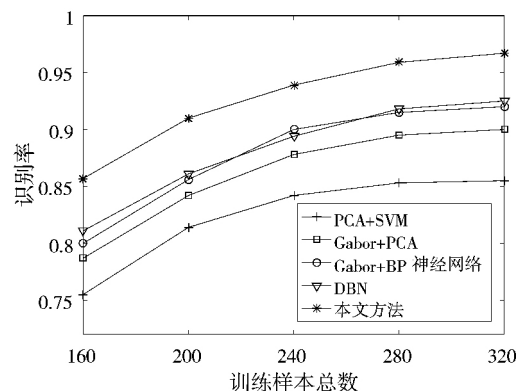


图7 UMIST 数据库上识别率对比

### 3.3 Yale-B 人脸库实验

Yale-B 标准人脸库包含 38 人共 21888 幅图像, 每人有 64 种光照、9 种姿态。考虑到实际应用中光照角度过大并不适用, 本实验将光照角度大于 85 的人脸图像去除, 每人保留 35 幅图像, 从中随机选取 26 幅图像作为训练样本, 其余作为测试样本。本实验采取与 ORL 人脸库上相同的实验策略进行参数配置。

表 1 Yale-B 人脸库上识别率对比

方法	识别率/%
PCA + SVM	62.75
Gabor + PCA	87.76
Gabor + BP 神经网络	92.08
DBN	92.34
本文方法	96.13

由表 1 中实验结果可知, 在多角度光照的人脸图像中, 本文所提方法与其他 4 种方法相比识别率最高, 达到了 96.13%, 进一步表明本文方法不仅能够学习到人脸图像的抽象特征, 而且对光照变化具有很好的稳健性。

## 4 结束语

本文提出一种结合 Gabor 小波与深度学习的人脸识别方法, 利用 Gabor 小波对光照、姿态等变化具有鲁棒性这一优点, 通过 Gabor 小波变换获取不同尺度和方向的人脸 Gabor 特征, 经下采样和 RBM 对 Gabor 特征进行有效降维, 结合 DBN 进行训练及识别。通过 ORL、UMIST 和 Yale-B 人脸库上的对比实验可知, 本文所提方法识别性能优于其他现有方法, 能够有效地提取人脸图像的抽象特征, 且很好地降低了光照、姿态等因素对识别效果的影响, 最终能够实现对人脸图像的准确识别。

#### 参考文献:

- [1] KIRBY M, SIROVICH L. Application of the Karhunen-Loeve procedure for the characterization of human faces [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1990, 12(1): 103-108.
- [2] SHAKUNAGA T, SHIGENARI K. Decomposed eigenface for face recognition under various lighting conditions [C]// Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2001, 1: 864-871.
- [3] BELHUMEUR P N, HESPAHNA J P, KRIEGMAN D J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1997, 19(7): 711-720.
- [4] PANG Y, YUAN Y, LI X. Gabor-based region covariance matrices for face recognition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(7): 989-993.
- [5] 许亚军, 李玮欣. 基于 Gabor 小波变换和神经网络的人脸识别研究 [J]. 中国电子科学研究院学报, 2017, 12(5): 534-539.
- [6] 朱树先, 张仁杰. BP 和 RBF 神经网络在人脸识别中的比较 [J]. 仪器仪表学报, 2007, 28(2): 375-379.
- [7] 陈海霞, 崔茜. 基于 Gabor 小波和 PCA 的人脸识别 [J]. 东北大学学报(自然科学版), 2014, 46(4): 77-80.
- [8] BENGIO Y. Learning deep architectures for AI [J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127.
- [9] BRAVERMAN M. Poly-logarithmic independence fools bounded-depth Boolean circuits [J]. Communications of the ACM, 2011, 54(4): 108-115.
- [10] BENGIO Y, LECUN Y. Scaling learning algorithms towards AI [J]. Large-Scale Kernel Machines, 2007(1): 321-360.
- [11] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [12] 张雯, 王文伟. 基于局部二值模式和深度学习的人脸识别 [J]. 计算机应用, 2015, 35(5): 1474-1478.
- [13] LYONS M, AKAMATSU S, KAMACHI M, et al. Coding-facial expressions with Gabor wavelets [C]// IEEE International Conference on Automatic Face & Gesture Recognition. IEEE Computer Society, 1998: 200-205.
- [14] VUKADINOVIC D, PANTIC M. Fully automatic facial feature point detection using Gabor feature based boosted classifiers [C]// IEEE International Conference on Systems, Man and Cybernetics. IEEE, 2005: 1692-1698.
- [15] 孔锐, 张冰. Gabor 滤波器参数设置研究 [J]. 控制与决策, 2012, 27(8): 1277-1280.
- [16] 徐永红, 侯景, 赵艳茹, 等. Gabor 小波幅值和相位特征人脸识别方法比较 [J]. 计算机工程与应用, 2012, 48(15): 195-200.
- [17] LIU C, WECHSLER H. Independent component analysis of Gabor features for face recognition [J]. IEEE Transactions on Neural Networks, 2003, 14(4): 919-928.
- [18] VENKATANARAYANA K, MANOJ V V R, SWATHI K S. Enhanced face recognition based on PCA and SVM [J]. International Journal of Computer Applications, 2015, 117(2): 40-42.